

Commission sur
l'avenir des soins de santé
au Canada



Commission on the
Future of Health Care
in Canada

ÉTUDE N° 35

L'influence des modes de rémunération des médecins sur l'efficiency du système de soins

par

Michel Grignon, Valérie Paris et Dominique Polton
avec la collaboration de **Agnès Couffinhall et**
Bertrand Pierrard
CREDES, Paris (France)



novembre 2002

N° de Catalogue CP32-79/35-2002F-IN

ISBN 0-662-87884-1

Bien que les opinions exprimées dans ces études n'engagent que les auteurs, chaque étude a été soumise à un processus indépendant d'examen par des pairs. La Commission tient à remercier l'Institut des services et de politiques de la santé (ISPS), des Instituts de recherche en santé du Canada, pour avoir assuré la surveillance et la bonne marche de ce processus d'examen. La participation des auteurs, des examinateurs et de l'ISPS à la réalisation de ces études constitue un apport important aux travaux de la Commission et contribuera à l'héritage qu'elle laissera.

Le générique masculin est utilisé uniquement dans le but d'alléger le texte.

Table des matières

Points saillants	iv
Sommaire	v
Liste des sigles et signification	viii
Question générale	1
Introduction	1
Les différents mécanismes de rémunération des médecins	3
Partie 1 : La littérature économique	5
Les faits et les attendus théoriques simples	5
L'arbitrage coût-qualité-équité : position du problème	6
L'arbitrage coût-qualité-équité : le pouvoir discrétionnaire du médecin	7
L'arbitrage coût-qualité-équité : la capacité du médecin d'augmenter la prescription	13
Partie 2 : À la recherche du mode de rémunération optimal : les enseignements d'une variété d'expériences	14
Combinaisons d'incitations au sein de la rémunération individuelle des médecins	14
Combinaisons d'incitations différentes au niveau individuel et collectif	18
Le groupe de médecins comme lieu d'intégration des objectifs cliniques et économiques	20
Incitations non financières	22
Conclusion : Quels enseignements pour le contexte canadien?	23
Questions spécifiques	24
Notes	27
Bibliographie	28

Points saillants

Cette étude traite de l'impact des modes de rémunération des médecins sur les coûts des soins médicaux.

Il existe trois grands modes de rémunération des professionnels qui dispensent des soins : le paiement à l'acte, la capitation et le salariat. En théorie, le paiement à l'acte peut inciter le professionnel à produire un volume de soins trop important, et les modes forfaitisés peuvent mettre la qualité et l'égalité d'accès en danger. L'importance de ces risques respectifs dépend fondamentalement de la nature et de l'étendue du pouvoir discrétionnaire du médecin sur le patient. Si le médecin influence simplement la demande, les risques sont limités par le contrôle exercé par le patient; si le médecin induit la demande, les risques sont plus importants.

Il n'existe pas de test empirique satisfaisant de la nature du pouvoir discrétionnaire du médecin. Pour autant, les études expérimentales comparant le volume et la qualité des soins selon le mode de rémunération concluent majoritairement à un effet important du paiement à l'acte sur les volumes produits par les médecins de première ligne (généralistes ou pédiatres). Réciproquement, la capitation augmente l'adressage et le volume prescrit, ce qui peut déporter les coûts sur d'autres postes de dépenses de santé. Des expériences naturelles dans le contexte norvégien montrent enfin que l'impact du mode de rémunération est plus faible que celui des caractéristiques sanitaires du patient.

Dans le monde réel (hors des expériences naturelles), les modes de paiement sont rarement purs :

- Première approche : moduler le paiement à l'acte par des incitations financières. Au Royaume-Uni, la délégation de budgets pharmaceutiques aux médecins généralistes a permis de réduire la consommation de médicaments. Ces incitations financières font cependant craindre des conflits d'intérêts entre le revenu du médecin et la qualité qu'il doit fournir. D'autres incitations sont fondées sur des performances sanitaires à atteindre, notamment en matière de dépistage ou de vaccination. Les études sont contradictoires, mais la majorité indiquent un impact positif significatif de ces incitations.
- Deuxième approche : faire dépendre le niveau de la rémunération d'un résultat collectif. Ce type d'approche semble inefficace si la collectivité visée est trop vaste (ensemble des médecins d'un pays), mais le paiement à la capitation de groupes de médecins semble une voie prometteuse pour concilier la maîtrise des coûts et l'organisation efficiente des soins.
- Troisième approche : les incitations non financières (protocoles, évaluation des pratiques). Peu d'études empiriques ont été réalisées sur ce type d'outils incitatifs, mais les premiers résultats indiquent un impact faible, sauf si les « messages » sont communiqués par des leaders d'opinion.

Sommaire

Une littérature économique abondante tente de modéliser la prestation et la consommation des soins comme un marché caractérisé par de nombreuses asymétries d'information, sur lequel le profit est une motivation possible du médecin.

Dans cette littérature, on s'intéresse à ce qui traite de l'impact des modes de rémunération du médecin sur le volume de services qu'il fournit ou prescrit au patient, la qualité de ces services et les inégalités d'accès aux soins.

On distingue trois grands modes de rémunération des médecins : paiement à l'acte, capitation (paiement forfaitaire par patient pris en charge pendant une période de temps donnée), et salariat. Il s'agit bien sûr d'une typologie simplificatrice, la plupart des systèmes nationaux de santé employant une combinaison de ces trois modes de paiement. En outre, le paiement dit à l'acte résulte lui aussi d'une certaine forfaitisation de l'activité médicale.

Les attendus théoriques sont les suivants : en mode de paiement à l'acte, le prix que reçoit le médecin tient compte des coûts fixes de son activité ainsi que des coûts variables, ce qui fait que l'acte marginal supplémentaire accroît toujours le profit monétaire du médecin. Le paiement à l'acte crée donc les conditions favorables à une quantité de services plus importante que ce qu'aurait souhaité le patient; pour peu que le médecin puisse influencer la demande du patient, le paiement à l'acte entraînera une inflation d'actes médicaux. En revanche, en mode capitation, le médecin n'a aucune incitation à produire trop d'actes; on craint au contraire qu'il ne restreigne son effort en réduisant l'accès du patient aux soins, notamment aux ressources immatérielles nécessaires à la production de soins, comme le temps consacré au patient, la capacité et la volonté de l'écouter, et l'effort mental consacré au cas, ce qu'on appelle souvent la qualité des soins. Enfin, toute forfaitisation fait courir un risque de sélection des cas : le médecin percevant la même somme pour des cas nécessitant un effort différent privilégiera les cas les plus simples.

Toujours du côté de la théorie, ces attendus se vérifient plus ou moins selon la nature de l'influence que peut exercer le médecin sur la consommation de soins des patients : une influence forte renforce les effets négatifs attendus (inflation des volumes en paiement à l'acte, qualité inférieure en capitation), une capacité de contrôle de la part du patient les atténue. Deux modèles de l'interaction médecin-patient rendent compte de ce degré d'influence : la concurrence imparfaite et l'induction. En concurrence imparfaite, le patient est à même de juger l'action de son médecin, mais il ne sanctionnera pas les comportements qu'il juge négativement aussi brutalement qu'il le ferait en concurrence parfaite : cette imperfection de concurrence vient de l'attachement du patient à son médecin traitant, qui le connaît personnellement. Dans l'induction, le patient subit plus fortement l'influence du médecin et ne peut contrôler *ex post* le comportement du médecin : ce dernier est en mesure de persuader le patient qu'il reçoit exactement les soins qu'il désire et il n'y a alors plus aucune concurrence.

Le problème est qu'il n'existe aucun test empirique satisfaisant permettant de départager ces deux modèles théoriques. On se tourne alors vers des observations dans le monde réel des interactions patient-médecin sous différents modes de rémunération. Pour inférer de telles observations sur l'impact des modes de rémunération, il faut être certain que l'on compare toutes choses égales par ailleurs. Deux expériences naturelles (affectation au hasard de médecins d'une clinique à des modes de paiement) confirment que le paiement à l'acte augmente les volumes, mais aussi que le niveau de rémunération semble jouer plus que le mode de rémunération. Les résultats des comparaisons « avant-après » (changement brutal et exogène du mode de paiement) montrent que l'impact du mode de rémunération dépend fortement de l'environnement du médecin, notamment de la pression des pairs. Un autre résultat est que le paiement à la capitation augmente la prescription de médicaments.

Dans le monde réel, il est rare qu'un mode de paiement soit pur : la plupart des payeurs cherchent une combinaison optimale, s'appuyant sur l'idée que le degré d'influence du médecin dépend du type de soins fournis. L'innovation majeure a consisté à introduire, au sein de paiements classiques, des correctifs à la marge en fonction de certains résultats sanitaires, le plus souvent mesuré sur une population large (taux de vaccination), ou financiers; on parle ici de paiement à la performance. Ainsi, au Royaume-Uni, la part du revenu indépendante du niveau de services représente 65 %, les paiements à l'acte 25 % et les paiements à la performance 10 %.

Les incitations aux résultats financiers sont efficaces, notamment pour réduire les taux d'hospitalisation, mais ils posent un problème de conflit d'intérêts de nature à abaisser le degré de confiance des patients pour leur médecin et à changer la nature de l'interaction patient-médecin.

Les payeurs préfèrent aujourd'hui les incitations aux résultats sanitaires, plus compatibles avec l'éthique médicale. Au Royaume-Uni, les *target payments* semblent avoir un impact positif sur les taux de vaccination ou de dépistage (observation simple avant-après). Les résultats d'études randomisées sont contradictoires : aucun impact sur le dépistage du cancer d'assurés Medicaid, mais impact positif sur la vaccination d'assurés Medicare ou sur des indicateurs de qualité comme le taux de césarienne, la durée de séjour hospitalier ou le taux de certification, pour des assurés du secteur privé. Rares sont les études qui analysent à la fois l'impact sur les coûts et la qualité.

Une autre innovation consiste à compenser les incitations individuelles par des incitations collectives : en mode paiement à l'acte, la valeur unitaire diminue quand le volume collectif augmente. Appliqué à un collectif trop large, ce système engendre un effet de passager clandestin, chaque individu augmentant ses volumes pour compenser la baisse de prix. Des contrôles individuels d'activité, ou des tarifs dégressifs sont alors nécessaires. Une autre solution consiste à réduire la taille du collectif bénéficiaire des incitations, sur le modèle des *primary care groups* britanniques, ou des *medical groups* aux États-Unis; la difficulté est alors de trouver l'équilibre entre la responsabilisation des médecins pour le collectif, et la solidité financière et technique du groupe, qui suppose une taille minimale.

Enfin, la plupart de ces incitations collectives reposent sur des éléments non financiers, comme la diffusion de protocoles de bonnes pratiques, ou l'évaluation de la pratique individuelle par les pairs, voire le conventionnement sélectif sur la base de pratiques définies en commun. L'impact de ces incitations non financières n'a cependant pas encore été évalué.

Liste des sigles et signification

- HMO :** *Health Maintenance Organization*
Régime de soins de santé à panel de médecins restreint (le patient n'est pas pris en charge s'il consulte hors panel)
- MEDICAID :** Assurance-maladie publique pour les pauvres aux États-Unis
- MEDICARE :** Assurance-maladie publique pour les plus de 65 ans et les personnes invalides aux États-Unis
- PCG/T :** *Primary Care Group* ou *Trust*
Groupe de médecins généralistes au Royaume-Uni, délégataire d'un budget global et prenant en charge tous les soins d'une population
- PPA :** Parité de pouvoir d'achat
Unité permettant de comparer des dépenses d'un pays à l'autre en s'affranchissant de la variabilité des taux de change
- PPO :** *Preferred Provider Organization*
Régime de soins de santé à panel de médecins recommandés (le patient n'est pris que partiellement en charge s'il consulte hors panel)

Question générale

Comment les mécanismes de rémunération et les incitations influent-ils sur l'utilisation et les coûts des services liés aux soins de santé? Dans quelle mesure des changements au niveau des politiques, des pratiques et des incitations changeraient-ils les structures de coût et le cheminement des soins de santé?

Introduction

L'efficacité du système de santé intéresse l'économie publique et la politique économique parce que la dépense de santé est toujours au moins mutualisée, voire socialisée. La question de l'efficacité a longtemps été limitée à la maîtrise de l'évolution des coûts des soins de santé, mais elle se pose maintenant de façon plus complexe : comment concilier une dépense raisonnable avec une exigence de qualité et d'équité?

Cette question a motivé une abondante littérature économique, modélisant les caractéristiques du marché des soins médicaux et les comportements des producteurs et des consommateurs de soins, afin de proposer des contrats optimaux au régulateur pour satisfaire ces trois objectifs contradictoires (maîtrise de la dépense, qualité, équité). Cette littérature théorique, centrée sur les comportements cachés au sein du trio payeur-patient-médecin, a débouché sur des questions empiriques non résolues, notamment au sujet de l'étendue de l'influence du médecin sur le patient.

Nous nous intéressons ici au volet de cette littérature qui traite de l'impact du mode de rémunération du médecin sur le volume de services qu'il fournit ou prescrit au patient, la qualité de ces services, et les inégalités d'accès aux soins. La façon de payer le médecin est un vrai choix stratégique et il n'existe pas de mode de paiement « naturel ». En effet, parce qu'il est soumis à une obligation de moyens et non de résultats (pour des raisons d'incertitude), le médecin n'est pas véritablement un producteur classique, qui vend un résultat et prend le risque que celui-ci ait une valeur inférieure à son coût de production. On parle du reste de rémunération des médecins, ce qui en fait des facteurs de production, plutôt que de prix de vente des services médicaux. Pour autant, le médecin n'est pas non plus un simple facteur de production, parce que son degré d'autonomie dans l'agencement des actions élémentaires qui concourent à la production de son service est une composante de la qualité des services qu'il offre.

Après une brève présentation des différents modes de rémunération des médecins, nous synthétisons, dans la partie 1, la littérature économique consacrée à leur impact sur le coût des soins de santé. Nous explicitons les enjeux théoriques des différents mécanismes de paiement et examinons dans quelle mesure les études empiriques menées dans différents contextes nationaux valident ces attendus théoriques.

Les économistes font l'hypothèse que, même si le profit du médecin n'est pas l'unique moteur de la décision médicale, il en est un élément. Cette hypothèse est partagée et confortée par la sociologie médicale, qui montre que, lorsque les conséquences d'une décision médicale sur

le bien-être du patient sont incertaines, les arbitrages du médecin en termes de loisir, de réputation ou de revenu prennent une place plus importante. Il n'est alors plus sûr que le patient reçoive exactement la quantité de soins nécessaire à son traitement.

Implicitement, les questions d'analyse économique sont donc du type suivant : étant données les motivations des médecins, quel impact a le mode de paiement sur la dépense médicale? Cette attitude ne signifie pas que les motivations autres que le profit (les normes professionnelles, le respect des pairs ou l'intérêt du patient) sont sans importance, mais plutôt qu'on peut faire l'hypothèse que les modes de rémunération sont sans impact sur elles. Dans la réalité, cette hypothèse est certainement fautive et les expériences de modification des modes de rémunération des médecins permettent alors de comprendre pourquoi les mécanismes mis en place peuvent s'écarter fortement des recommandations économiques, notamment parce que le mode de rémunération peut interférer avec ces autres motivations du médecin (partie 2).

Les différents mécanismes de rémunération des médecins

Il est d'usage de distinguer trois grands modes de rémunération des médecins :

- La rémunération à l'acte
- La capitation, qui est un paiement forfaitaire par patient pris en charge pendant une période de temps donnée
- Le salariat

Selon les pays, l'une ou l'autre de ces trois formes de rémunération est utilisée pour les généralistes. Les spécialistes sont en général payés à l'acte ou salariés.

Cette typologie classique suscite quelques réflexions.

Tout d'abord, la rémunération globale d'un médecin peut combiner plusieurs méthodes de paiements : les paiements mixtes sont pratiqués dans certains systèmes de santé, comme on le verra dans la partie 2.

Ensuite, même s'il est commode d'opposer ces trois mécanismes, on peut aussi les considérer comme s'échelonnant sur une échelle de forfaitisation croissante, qui permet aussi d'imaginer des situations intermédiaires :

- À une extrémité, le paiement à l'acte correspond au paiement le plus décomposé, le moins forfaitaire, de l'activité médicale. Malgré tout, il s'agit encore d'un forfait : lorsqu'une consultation est payée à un tarif unique, comme en France par exemple, il importe peu que cette consultation soit courte ou longue, qu'un examen clinique soit réalisé ou que le médecin renouvelle simplement une ordonnance. On peut d'ailleurs noter que les degrés de forfaitisation des nomenclatures utilisées dans les différents pays pour tarifier les actes varient et qu'une analyse comparative de ces nomenclatures pourrait permettre de qualifier différents types de paiements à l'acte.
- La capitation représente un degré de forfaitisation supplémentaire, puisque le médecin est payé de la même manière pour un patient quelle que soit la quantité de services qu'il lui fournit. Le revenu du médecin ne varie qu'en fonction du nombre de malades pris en charge.
- Le salariat représente la situation extrême du paiement forfaitaire dans laquelle le revenu est invariant quelle que soit l'activité du médecin, qu'il s'agisse du nombre de patients soignés ou de la quantité de services fournis par patient.

Sur cette échelle, on peut imaginer des situations intermédiaires, par exemple entre le paiement à l'acte et la capitation globale : un paiement forfaitaire par épisode de soins ou un paiement forfaitaire pour le suivi d'une pathologie chronique pendant une période de temps. Quelques expérimentations avec de tels mécanismes seront évoquées dans la partie 2.

En résumé donc, même si l'on oppose classiquement – et c'est ce que nous ferons dans la revue de la littérature économique ci-dessous – trois modes de rémunération « archétypaux », il n'est pas interdit de concevoir des modalités intermédiaires.

Enfin, dans les formes de paiement des professionnels de la santé, on voit apparaître des mécanismes qui ne correspondent à aucun des trois modèles évoqués ci-dessus, mais qui s'apparentent à des mécanismes d'intéressement aux résultats, certains financiers (contrôle des coûts) et d'autres visant la qualité du service médical.

Ce type de mécanismes, dont des illustrations seront données dans la partie 2, pourrait être considéré comme un quatrième mode de rémunération, alternatif aux trois précédents. C'est d'ailleurs ce que font certains auteurs qui distinguent désormais les quatre modes, donnant souvent au quatrième l'appellation de « paiement à la performance ». Dans la pratique, cependant, ce type d'intéressement n'intervient que comme un complément à la marge du mode principal de rémunération et peut difficilement être considéré comme une alternative; il s'agit plutôt d'un appoint aux trois grands modes de paiement. Dans l'avenir, il est possible qu'il prenne une part croissante dans la rémunération.

Partie 1 : La littérature économique

Voici un bref aperçu de la littérature économique traitant de l'impact du mode de paiement des services des médecins sur le coût budgétaire total de la santé.

Les faits et les attendus théoriques simples

Sur la base des données agrégées, la dépense totale de santé d'un pays est affectée par les variables d'offre en général et par le mode de rémunération des médecins en particulier. Sur dix-neuf pays de l'OCDE, Gerdtham et coll. (1992) constatent que, toutes choses égales par ailleurs, le « paiement à l'acte des médecins » (c'est-à-dire une variable indiquant que ce mode de paiement est majoritaire dans le pays) augmenterait de 11% la dépense de santé. Une étude plus récente, utilisant des données de panels (22 pays de l'OCDE entre 1970 et 1991), ne trouve plus d'effet propre au paiement à l'acte, mais un effet indirect : la densité médicale augmente la dépense quand les médecins sont rémunérés à l'acte (Gerdtham et coll. 1998). Ces liens au niveau des données agrégées sont cependant difficiles à interpréter car il est possible que la dépense nationale totale et des arrangements institutionnels particuliers, comme le paiement à l'acte, soient tous deux corrélés positivement à une troisième variable, par exemple le fait que les médecins exercent de manière isolée.

La littérature économique tente donc de décrire le cheminement micro-économique par lequel le mode de rémunération du médecin pourrait avoir une influence sur la dépense médicale; pour formaliser ce cheminement, on caractérise un mode de paiement par le degré auquel le paiement est forfaitisé. Dans la suite de cette première partie, nous nous concentrons sur l'opposition entre le système de paiement forfaitaire (capitation et salariat) et celui du paiement à l'acte.

Nous pouvons relever tout d'abord les attendus de l'impact des modes de paiement sur le coût des soins de santé en faisant abstraction des autres objectifs du système de santé (la qualité et l'équité). Pour ce faire, décrivons rapidement l'activité du médecin : pour soigner les patients, il utilise des ressources. Certaines ressources comme le cabinet doivent être présentes qu'il y ait ou non des patients; d'autres ne sont mobilisées qu'en présence d'un client. Dans la réalité actuelle, le médecin ne fixe pas lui-même son prix, il le reçoit d'un payeur. Les ressources utilisées dans la production sont de deux types, ressources matérielles quantifiables et ressources immatérielles difficilement quantifiables. Les premières sont, par exemple, les biens consommables de l'activité médicale, les secondes réunissent le temps que le médecin consacre à chaque patient, sa compétence (qui est fonction de sa formation continue, du soin qu'il prend à se tenir informé), sa capacité à écouter le patient, et son effort mental. L'économie désigne habituellement le niveau de ressources immatérielles par « l'effort du médecin ». Une différence importante entre ressources matérielles et immatérielles est que le payeur peut tenir compte des premières dans son paiement, mais plus difficilement des secondes. En effet, il est difficile pour un payeur de rémunérer une ressource quand il ne peut avoir la preuve qu'elle a bien été utilisée.

Muni de ce modèle simplificateur, que peut-on dire du mode de paiement?

En mode de paiement à l'acte, le prix tient compte des coûts fixes et des coûts variables (sinon, le médecin refuse de travailler) et est donc supérieur au coût moyen des seules ressources variables : soigner un patient supplémentaire apporte donc un profit financier au médecin. Si, de surcroît, le médecin dispose d'un pouvoir quelconque sur la demande du patient, le paiement à l'acte crée les conditions favorables à une quantité (et donc à une dépense) de services plus importante que ce qu'aurait souhaité le patient, ou que ce qui aurait été médicalement nécessaire dans son cas. Dans ce modèle, et toujours si le médecin exerce un pouvoir sur le patient, si le payeur diminue le prix unitaire, le médecin réagira en augmentant la quantité¹ pour maintenir son profit; on explique ainsi qu'une baisse du prix des soins entraîne une augmentation de l'offre de services.

En revanche, en mode capitation, le médecin n'est pas remboursé proportionnellement aux ressources qu'il engage et n'a donc aucun intérêt à la multiplication des actes, puisqu'il perd alors de l'argent ou gaspille son effort. Il préfère traiter le patient en utilisant le moins possible de ressources pour atteindre une qualité donnée.

L'arbitrage coût-qualité-équité : position du problème

Si la question se résumait à la maîtrise des coûts, on pourrait se demander pourquoi tous les régulateurs n'ont pas opté pour une forme ou une autre de paiement forfaitaire. Il existe bien évidemment des raisons historiques qui maintiennent le paiement à l'acte en dépit de la supériorité évidente du forfait comme instrument de maîtrise budgétaire (voir, par exemple, Emery, Auld et Lu 1999). Cependant, sur le territoire culturellement homogène des États-Unis, on constate des arrangements différents selon les payeurs : certains préfèrent des formes de capitation, d'autres le paiement à l'acte, ces derniers survivant et semblant même regagner des parts de marché au cours de la période récente.

Il existe donc certainement des raisons liées à l'efficacité économique pour expliquer que le paiement forfaitaire n'ait pas été adopté universellement :

- Une première raison est liée à la qualité du service médical obtenu : si la forfaitisation permet de résoudre, par construction, le problème de l'augmentation des volumes de soins dispensés par le médecin, elle fait peser le risque que le médecin fournisse trop peu de services (consultations, examens) et trop peu de ressources immatérielles (consultations plus courtes, moins d'attention) au patient. Le régulateur peut bien évidemment imposer un minimum de services par patient, mais il aura du mal à contrôler le niveau des ressources immatérielles. Si le médecin est payé pour la prise en charge d'un patient et n'est sanctionné d'aucune manière², il a tout intérêt à percevoir le forfait et à ne pas fournir les services et l'effort correspondants. La littérature économique tente d'éclaircir ce point et de définir les circonstances dans lesquelles un forfait met la qualité médicale en danger. Plus généralement, le forfait peut être vu comme une délégation au médecin de

la responsabilité d'agencer au mieux les services qui concourent à la santé, alors que le paiement à l'acte partage ce rôle entre médecin et régulateur. En présence de payeurs multiples, le paiement à l'acte pose un problème supplémentaire de non-cohérence des systèmes de prix, mais ce problème n'a été étudié à notre connaissance que pour l'offre hospitalière (Tai-Seale, Rice et Stearns 1998).

- Ce problème de qualité peut aussi avoir des conséquences sur la dépense : le médecin n'est pas seulement un producteur de services, il est aussi prescripteur. En règle générale, la capitation s'accompagne d'une responsabilité d'aiguillage du médecin de première ligne : il prend les décisions concernant les volumes d'analyses, de médicaments, de consultations spécialisées ou de soins hospitaliers. Si le médecin réduit l'intensité de son effort à cause du forfait, tout en restant soucieux de la santé de son patient, il peut fort bien générer un fort volume de coût de soins de santé en augmentant sa prescription.
- Enfin, se pose le problème de l'universalité de l'accès aux soins : le forfait entraîne toujours un risque financier pour le médecin; si les médecins ont de l'aversion pour les situations incertaines, ils vont refuser de prendre en charge les patients à forte variance de coût *ex ante*. Si les médecins savent mieux sélectionner que les payeurs, ils vont gagner de l'argent en traitant les personnes dont le forfait est élevé et en refusant les soins aux autres (voir, par exemple, Mulligan 2002).

Au total, la question devient donc : quelle forme de paiement permet le mieux de concilier les objectifs de maîtrise des dépenses, de qualité du service médical et d'universalité d'accès aux soins? Nous nous intéressons d'abord aux modèles théoriques qui expliquent le pouvoir discrétionnaire du médecin sur la consommation du malade et qui s'opposent sur l'existence d'une capacité du médecin à « induire » la demande de son patient (Evans 1974). Cette capacité à induire a des conséquences importantes pour notre propos, car elle augmente le risque que les paiements forfaitaires conduisent à une diminution du service médical reçu par le patient. Nous complétons notre présentation par une analyse du comportement du médecin comme prescripteur de services médicaux qu'il ne fournit pas lui-même.

L'arbitrage coût-qualité-équité : le pouvoir discrétionnaire du médecin

Approche théorique

Nous avons vu précédemment que le médecin tirait profit du paiement à l'acte s'il avait la possibilité d'imposer au patient plus de services que celui-ci n'en souhaite ou n'en a besoin. D'où vient une telle possibilité?

Un accord assez large semble se faire sur l'idée que le médecin dispose effectivement, dans l'interaction avec ses patients, d'une supériorité qui lui permet d'influencer la consommation de ses services.

Une raison pour laquelle le médecin dispose d'une supériorité est qu'il en sait plus que son patient. Une abondante littérature théorique traite du fait que la supériorité des connaissances suffit à un expert pour imposer une relation d'agence, c'est-à-dire une relation dans laquelle le profane s'en remet totalement à l'expert. Cette littérature conclut qu'en situation de concurrence des experts, et si le profane peut se faire une idée assez précise du lien existant entre l'effort de l'expert et son résultat, alors la supériorité des connaissances ne suffit pas à créer l'agence. L'expert vend un bien dit « d'expérience », sur lequel l'acheteur peut finir par se faire une idée. En revanche, si une incertitude radicale interdit à l'acheteur de se faire une idée quelconque sur l'impact de l'effort de l'expert sur le résultat qu'il obtient, alors la relation d'agence peut exister. On dit que l'expert vend un bien de « confiance » car l'acheteur profane doit lui faire une entièrement confiance.

Qu'en est-il pour les biens médicaux? Rice (1993) avance l'idée radicale qu'il n'existe simplement pas de demande de soins éclairée et autonome de la part du patient. Il appuie sa démonstration sur le fait que, dans l'expérience de la Rand, les soins supplémentaires que s'offrent les patients les mieux couverts ne sont pas, au dire d'experts indépendants, des soins vraiment appropriés compte tenu de leur état de santé.

L'idée qu'il n'existe aucune demande autonome du patient en matière de soins est sans doute excessive (nous faisons tous des choix sanitaires indépendants de nos médecins) mais, en revanche, il est certain que le service médical rendu par le médecin est pour une part non observable. Cette difficulté à juger le service médical rendu constitue sans doute l'une des raisons pour lesquelles les médecins n'ont pas le droit, sauf exception, dans les pays développés, de détenir des pharmacies ou d'exercer le métier de vendeur de médicaments : s'ils pouvaient augmenter leur profit de cette manière, ils pourraient abuser de leur position d'expert pour convaincre les patients de leur intérêt à consommer des médicaments.

On peut aussi faire quelques prédictions *a priori*. Par exemple, il est certainement plus difficile d'observer la qualité du service d'un spécialiste que l'on voit rarement : pour résoudre l'asymétrie d'information, le régulateur préférera faire observer la qualité du service de ce spécialiste par un autre médecin, en soumettant le recours à un adressage³ par un médecin traitant de première ligne. En revanche, les preuves empiriques d'un degré d'induction de la part de médecins sont très difficiles à réunir : même si l'on mesure une différence de traitement pour deux malades identiques auprès de deux médecins différents, et si l'on arrive à relier cette variation de pratique à l'intérêt économique du médecin, encore faudra-t-il être certain que la variation n'est pas ressentie par le patient comme une dégradation du service qu'il reçoit.

En effet, la relation d'agence suppose non seulement que le médecin impose ses vues au patient, mais aussi que le patient ne s'en aperçoive pas (il fait confiance à l'expert). Or, comme le montre McGuire (2000), on peut très bien imaginer que le médecin parvienne à imposer ses vues même si le patient s'aperçoit que le médecin ne fait pas exactement ce que le patient aurait désiré, ou ce que le médecin ferait pour lui-même. En termes techniques, la relation d'agence n'est pas nécessaire pour qu'existe un pouvoir discrétionnaire du médecin sur le patient : il suffit que le médecin soit en situation de concurrence imparfaite face aux patients pour que ceux-ci acceptent de sa part une prestation de services non totalement satisfaisante.

Une bonne raison pour qu'il y ait situation de concurrence imparfaite des médecins face aux patients est l'attachement du patient à certaines caractéristiques intrinsèques de son médecin qui fait qu'il n'accepterait pas volontiers d'en chercher un autre. Cet attachement du patient à « son » médecin (médecin de famille, source habituelle de soins) est abondamment documenté et s'explique par le fait qu'une fréquentation de longue durée permet un suivi spécialisé du cas du patient. Cet attachement du patient au médecin concerne évidemment plus le médecin généraliste (ou de première ligne) ou encore le spécialiste pour une maladie chronique.

Cet attachement du patient à un médecin particulier fait que les médecins ne sont pas des substituts parfaits : chaque médecin dispose d'une latitude face à son patient, et peut lui imposer des choix différents de ceux que ferait le patient spontanément s'il pouvait sans dommage remplacer son médecin par un autre. Disposer d'une latitude se traduit en termes économiques par « disposer d'un pouvoir de marché » et signifie se comporter comme un monopole. Précisément, le médecin est en situation de concurrence monopoliste face à son patient : comme un monopole, le médecin peut tirer le meilleur parti de la disposition à payer de son patient en jouant sur les prix et les quantités offertes (alors qu'en concurrence parfaite, il y a un seul prix, celui du marché); pour autant, si le médecin « en fait trop » (s'éloigne trop des attentes ou des besoins de son patient), celui-ci le quitte pour un autre médecin.

On peut même souligner que l'attachement et l'agence opèrent de manière complémentaire : quand le patient voit rarement le médecin, il a peu d'attachement à la personne du médecin, mais il a aussi plus de chance d'être en relation d'agence et de devoir lui faire confiance aveuglément; en revanche, quand le patient est engagé dans une relation de long terme avec son médecin, il observe plus facilement la qualité du service, mais il a plus de difficulté à faire jouer la concurrence et à s'opposer à des prestations de services qui le laissent insatisfait.

Quelle conséquence peut avoir la nature de ce pouvoir discrétionnaire du médecin sur le problème qui nous occupe? Schématiquement, plus on s'approche d'une situation de concurrence monopoliste et plus le médecin doit tenir compte de la menace que fait peser sur lui le départ du patient : en mode paiement à l'acte, cette menace limite le pouvoir du médecin de dispenser trop d'actes dont le patient ne veut pas; en mode capitation, elle limite le pouvoir du médecin de trop réduire le volume de services et, surtout, l'effort qu'il consacre à son patient. Si l'attachement des patients pour un médecin donné est une variable continue, le médecin opère un arbitrage entre la taille de sa liste et le bénéfice qu'il retire de chaque inscrit. On dit alors que le médecin influence la consommation du patient.

En revanche, plus on s'approche de la relation d'agence, et moins le patient est apte à faire peser une menace sur le médecin, qui peut fournir trop de services s'il est rémunéré à l'acte, ou pas assez d'effort en mode de paiement forfaitaire. En ce cas, on parle d'induction de la demande par le médecin.

Une abondante littérature tente de prouver empiriquement et de manière indirecte que le médecin dispose d'un pouvoir discrétionnaire absolu, car non contrôlé par le patient. Un test classique consiste à mesurer la réponse des médecins à une variation de la densité médicale en contexte de paiement à l'acte. Rappelons que, en cas d'induction, le médecin est hors

concurrence : si la densité médicale augmente, le médecin aura un moins grand nombre de patients en moyenne (dans le cas de l'induction, la qualité du médecin joue peu sur son attractivité puisqu'il n'est pas en concurrence) et réagira donc pour préserver son revenu en augmentant le volume des services utilisés par ses patients. En revanche, en concurrence monopoliste, l'augmentation de la densité médicale renforce le risque pour le médecin que son patient le quitte (le coût d'une rupture est moindre pour le patient) et il ne va donc pas imposer un écart encore plus important entre la consommation souhaitée par le patient et les services fournis.

Outre les questions empiriques (caractère vraiment exogène de l'augmentation de densité, difficulté à observer des variations importantes), ce test n'est pas parfait. Il repose sur l'hypothèse contestable qu'en concurrence monopoliste, l'augmentation de la densité médicale facilite le départ du patient mécontent; en fait, si l'attachement du patient dépend de caractéristiques personnelles du médecin qui ne sont pas affectées par la densité médicale, on peut observer une augmentation du nombre d'actes par suite d'une augmentation de la densité médicale dans un contexte d'influence. Pauly et Satterthwaite (1981) font une objection du même type : une densité supérieure de médecins augmente la capacité de chaque médecin d'influencer la demande par un mécanisme de médicalisation des attentes de la population.

Au total, il semble qu'on ne puisse trancher sur le degré d'induction à partir de mesures du niveau d'activité par médecin. Les enquêtes directes auprès des patients sur le degré de satisfaction et sur les raisons pour lesquelles ils choisissent ou quittent tel ou tel médecin seraient sans doute d'un grand intérêt sur ce point. Mais force est de constater qu'il n'y a pas eu de mesure du degré d'induction fondée sur des déclarations directes de patients. Une indication précieuse ressort du fait qu'aux États-Unis, les assurés du programme Medicare les plus malades préfèrent rester dans les régimes de paiement à l'acte et garder la liberté d'accès au médecin de leur choix plutôt que de bénéficier de la couverture plus étendue proposée par les HMO. On peut donc craindre que la latitude du médecin à convaincre le patient que ce dont il a besoin est justement ce qui maximise le revenu du médecin n'augmente avec le degré de morbidité du patient.

Approches empiriques

Face à un constat de variation de la pratique médicale, il est impossible de savoir *a priori* si une pratique est excessive ou si c'est l'autre qui est restrictive. On peut comparer *a posteriori* les pratiques à une norme technique, comme une recommandation, ou à une norme statistique comme la pratique moyenne, mais rien ne valide de telles comparaisons.

On ne peut donc que mesurer l'impact du mode de rémunération du médecin sur le volume de services consommé par patient, la qualité et l'universalité, sans se prononcer à partir de telles études sur le bon type de paiement. Pour mesurer cet impact, nous cherchons des comparaisons toutes choses égales par ailleurs, sauf le mode de paiement. Scott et Hall (1995) ont recensé sept études empiriques publiées en anglais sur ce thème et soulignent que ce nombre limité est lié au fait qu'il s'agit d'études « opportunistes », utilisant des données recueillies à d'autres fins. Gosden et coll. (2001) recensent une huitième étude.

Pour que la comparaison ait un sens, il faut d'abord être certain que l'on se trouve dans des contextes comparables : comparer l'impact de systèmes de paiements différents dans des zones géographiquement ou socialement très éloignées n'aurait pas grand sens. Il faut donc trouver des contextes homogènes avec une variabilité suffisante des modes de paiement. Un deuxième problème lié à ce type de comparaison est qu'il est difficile de s'assurer que toutes choses autres que le mode de paiement soient bien égales par ailleurs : le fait qu'un patient opte pour un médecin rémunéré à l'acte ou en capitation peut être corrélé avec des caractéristiques non observables, mais elles-mêmes liées à l'utilisation du système de soins. De même, les médecins peuvent se spécialiser par types de contrats (Glied 1998), là encore pour des raisons non indépendantes de leur mode de travail et du volume de services qu'ils dispensent. Enfin, il est difficile de s'assurer que le niveau du paiement est bien le même d'un mode de paiement à l'autre et qu'on ne mesure pas l'effet combiné d'une différence de mode et de niveau de paiement.

Hickson et coll. (1987) ont mené une étude expérimentale en contrôlant les deux premiers problèmes : quinze pédiatres, membres d'une même clinique, sont alloués aléatoirement à un mode de paiement (salarial ou paiement à l'acte). L'étude compare alors le nombre de visites préventives des salariés et des médecins payés à l'acte : à caractéristiques observables des médecins et des patients contrôlées, les pédiatres payés à l'acte montrent significativement plus de séances.

En outre, ce surcroît de volume reflète une variation ambiguë de la qualité : les patients traités par les médecins payés à l'acte affirment bénéficier d'un meilleur suivi; en revanche, le volume des séances préventives qu'on leur propose s'éloigne plus des recommandations de l'Académie américaine des pédiatres que celui des médecins salariés. Gosden et coll. (2001) soulignent que cette différence moyenne entre médecins salariés et rémunérés à l'acte recouvre deux phénomènes distincts : les médecins salariés sont moins souvent en recommandation « excessive » de séances que ceux rémunérés à l'acte, mais les patients des médecins salariés observent aussi moins les recommandations de leur médecin.

Une autre étude expérimentale avec affectation aléatoire des médecins entre les modes de paiement est celle de Davidson et coll. (1992) qui porte sur 80 médecins traitant des enfants bénéficiaires du programme Medicaid aux États-Unis. L'étude suit trois groupes, le groupe de référence, payé à l'acte, un groupe payé à l'acte mais touchant un prix plus élevé, et un groupe payé à la capitation. Le volume de séances de soins primaires augmente pour les deux groupes « expérimentaux », ce qui souligne l'impact du niveau de rémunération, indépendamment du mode de paiement. Cependant, le volume augmente plus dans le groupe des médecins rémunérés à l'acte (+0.8-0.9) que dans le groupe des médecins rémunérés à la capitation (+0.5-0.6).

Le prix à payer pour ces gains méthodologiques (affectation aléatoire des médecins entre les modes de paiement) est évidemment la faiblesse de l'échantillon et le caractère peu généralisable des résultats.

Nous nous tournons donc vers des comparaisons « avant-après » dans lesquelles un même groupe de médecins et de patients s'est vu imposer par le payeur un changement de mode de paiement. En effet, nous pouvons faire l'hypothèse que, dans ce cas, le mode de paiement change vraiment de manière exogène et observer l'effet pur de tout biais. En outre, les réformes se font souvent autant que possible « à revenu constant » pour les médecins et nous pouvons donc mesurer l'effet du mode de paiement indépendamment de l'effet du niveau de revenu.

Une première observation de ce type a été faite par Stearns et coll. (1992) et révèle la faible efficacité de la capitation sur le volume de services fournis : dans un régime de soins de santé pour les fonctionnaires du Wisconsin, les médecins de première ligne sont passés du paiement à l'acte à la capitation individuelle. Simultanément, le tarif unitaire de l'acte de spécialiste est fortement réduit. Il y a eu paradoxalement augmentation du nombre de séances ambulatoires, mais c'est le nombre de séances chez les spécialistes qui a augmenté; il y a lieu de penser que les spécialistes ont accru leur auto-adressement pour compenser la baisse de tarif. Le nombre de séances chez les généralistes est resté inchangé et le passage à la capitation n'a donc pas eu d'effet. Il est toutefois possible qu'un autre élément ait perturbé l'expérience : simultanément, les généralistes ont aussi bénéficié d'un intéressement à limiter les dépenses hospitalières. Il est donc possible qu'ils aient été motivés par cet intéressement et aient mis davantage l'accent sur la médecine de première ligne pour éviter des hospitalisations. De fait, l'étude montre une baisse sensible du taux d'hospitalisation (mais non de la durée de séjour). Cette étude n'est pas totalement exempte de biais, car certains médecins de première ligne ont quitté le régime ou ont refusé des patients du régime (ou d'offrir certains services). Des patients ont pu aussi quitter le régime car, outre le changement de mode de rémunération, il y a eu passage en HMO exclusif (non prise en charge de l'assuré s'il consulte hors du panel), alors que le régime fonctionnait auparavant en PPO (moindre prise en charge de l'assuré s'il consulte hors du panel).

Une expérience contrôlant mieux ces départs volontaires a été menée au Danemark (Krasnik 1990). Les généralistes de la ville de Copenhague sont passés de la capitation pure à une formule combinée de capitation et de paiement à l'acte, tandis que les médecins de la région restaient en mode capitation. L'auteur peut donc comparer l'évolution des volumes dans le groupe « soumis à l'expérience » (les généralistes de Copenhague) et dans le groupe « témoin ». Le volume de services offert par les médecins du groupe « expérimental » a fortement augmenté à la suite du passage au paiement à l'acte.

Cependant, cet effet n'est pas universel et l'expérience norvégienne semble moins concluante quant à l'impact du paiement à l'acte. En utilisant des chroniques de contacts par patient et par médecin⁴ en Norvège, Kristiansen et Holtedahl (1993) montrent que le paiement à l'acte hausse le nombre de visites à domicile toutes choses égales par ailleurs (l'état de santé du patient étant cependant inconnu), mais que l'effet du mode de paiement reste faible par rapport à l'influence des caractéristiques individuelles du patient : le médecin reste plus motivé par le besoin sanitaire du patient que par son revenu⁵. Cette primauté du clinique sur l'économique en Norvège semble confirmée par Grytten et Sorensen (2001), qui montrent que les médecins rémunérés à l'acte ne font pas plus d'examen que les médecins salariés, même après avoir contrôlé les éléments qui incitent les médecins à choisir tel ou tel type de contrat. Selon ces auteurs, la primauté du facteur

clinique serait due principalement à l'existence d'un fort contrôle par les pairs en Norvège, qui empêcherait le médecin de tirer parti de sa supériorité sur le patient.

Outre un effet dommageable sur la qualité moyenne, l'induction peut aussi avoir une conséquence dommageable sur l'équité du système : les populations vulnérables (enfants, personnes âgées, malades mentaux, pauvres, minorités linguistiques) ont plus de mal que les autres à contrôler et à observer la qualité du service du médecin, et il est donc tentant pour celui-ci d'utiliser cette capacité à induire la demande principalement sur les populations vulnérables. Il est notable que les HMO rationnent la qualité clinique sur ces populations, tout en laissant la qualité moyenne inchangée (Miller et Luft 1997).

L'arbitrage coût-qualité-équité : la capacité du médecin d'augmenter la prescription

L'adressage et la prescription sont des éléments importants de la pratique du médecin et des coûts totaux des soins de santé. En France, 33 % des séances auprès d'un médecin seraient motivées par une recommandation d'un autre médecin. Au Royaume-Uni, 5 % des séances auprès d'un médecin de première ligne donnent lieu à un adressage (Coulter 1992) et on observe une très forte variation des pratiques des médecins, dont les caractéristiques des patients ne rendent pas compte (résumé par Scott 2000).

Une théorie simple, mais robuste, du comportement du médecin montre que la capitation et le salariat incitent le médecin à transférer à d'autres son propre travail et à augmenter le volume prescrit pour diminuer l'effort non rémunéré (temps passé avec le malade, écoute, auscultation).

Les données empiriques confirment largement cette prédiction. L'étude de Krasnik et coll. (1990) montre que le passage des généralistes au paiement à l'acte a significativement réduit leur taux d'adressage et de prescription. Rappelons que cette étude dispose d'un échantillon contrôlé dont le mode de paiement est resté inchangé au cours de la même période.

L'idée que le temps passé en consultation est plus court dans le mode de paiement à la capitation ressort des comparaisons internationales : d'après Sandier (1990), alors que le généraliste français, payé à l'acte, consacre quinze minutes (en 1979) en moyenne à chaque consultation, le généraliste anglais, payé à la capitation, se contente de huit minutes (en 1987). Cependant, à contexte homogène et pour des données individuelles, il n'est pas confirmé que la capitation ou le salariat diminue le temps de consultation : Kristiansen et Mooney (1993) trouvent même, avec des données norvégiennes, que les généralistes payés à l'acte passent moins de temps en consultation que les salariés, mais la différence est non significative.

Une solution pour éviter que la capitation n'augmente le coût des soins prescrits est de rendre le prescripteur responsable du budget total des soins pour une personne. Il s'agit d'une forme d'intégration verticale des soins, confiée au médecin de première ligne et que nous détaillerons ci-dessous.

Partie 2 : **À la recherche du mode de rémunération optimal :** **les enseignements d'une variété d'expériences**

La question des incitations liées aux modes de paiement a reçu une attention croissante depuis que les systèmes de santé font face à des enjeux majeurs d'arbitrage entre coût, qualité et équité. Réformes générales ou expérimentations locales, diverses solutions ont été mises à l'essai pour concilier ces objectifs potentiellement contradictoires au niveau des mécanismes de rémunération des professionnels. Nous les passerons en revue dans cette partie, en distinguant plusieurs approches.

Une première tente de combiner des incitations variées au sein même de la rémunération individuelle des médecins, avec des systèmes de paiement mixtes comprenant forfait, capitation et paiement à l'acte. Des incitations financières particulières y sont éventuellement incorporées; parmi celles-ci les formes les plus novatrices sont des mécanismes d'intéressement au résultat, financier ou sanitaire (paiement à la performance).

Une deuxième élabore des systèmes d'incitation financière différents au niveau individuel et au niveau collectif de la profession médicale.

Les conflits d'intérêts qui en résultent conduisent à envisager cette formule non plus au niveau d'un collectif très large de professionnels, mais au sein de groupes plus restreints : c'est le sens d'une troisième approche, qui se développe aujourd'hui, et qui confie à des groupes de médecins la conciliation d'objectifs financiers et de qualité.

Enfin, une dernière voie de réflexion dans la recherche du meilleur rapport coût/qualité consiste à associer aux incitations financières des incitations d'un autre ordre, comme l'information fournie au médecin, ou l'élaboration de références de pratique et de divers moyens de les diffuser. La dimension collégiale de l'exercice médical peut favoriser cette évolution des pratiques et l'on retrouve, par ce biais, un intérêt potentiel des groupes ou des réseaux évoqués ci-dessus.

Combinaisons d'incitations au sein de la rémunération individuelle des médecins

Avant d'aborder les systèmes de paiement mixtes, notons qu'aucun mode de paiement n'a émergé comme une alternative globale aux « grands standards » que constituent le paiement à l'acte, la capitation et le salaire.

Le paiement par épisode de soins, transposé du paiement des hôpitaux au groupe homogène de malades (DRG), a été considéré par le programme Medicare aux États-Unis au milieu des années 80, mais il ne semble pas avoir eu beaucoup d'applications concrètes en médecine ambulatoire. Il est envisageable dans le cas des soins techniques pour des épisodes aigus,

par exemple un chirurgien percevant une rémunération globale pour la phase pré-opératoire, l'acte et le suivi post-opératoire (Robinson 2001a); il a été étudié en cardiologie, par exemple pour le pontage coronarien (Showstack, Garnick et Rosenfeld 1987). Mais l'extension de ce mode de paiement à l'ensemble de l'activité médicale soulève de nombreux problèmes techniques.

En termes d'applications pratiques, l'innovation et l'inventivité se sont donc plutôt manifestées au niveau de la façon d'agencer différents éléments de paiement et d'aménager des incitations à la marge pour une petite fraction de la rémunération.

Un exemple intéressant de paiement mixte est celui des généralistes anglais, dont la rémunération comprend (Redbook 2002) :

- Une allocation de base destinée à prendre en charge une fraction des coûts de fonctionnement, sous forme d'un forfait par patient dégressif en fonction du nombre de patients. Une allocation supplémentaire compense les surcoûts liés à l'exercice en région rurale. Des majorations sont également versées en fonction de l'expérience acquise par le médecin (honoraires dits « de séniorité »).
- Une capitation par patient inscrit sur la liste du généraliste, avec trois niveaux selon l'âge (moins de 65 ans, 65 à 74 ans et 75 ans et plus); les patients des zones défavorisées donnent aussi lieu à une capitation additionnelle, reflétant leurs besoins plus importants.
- Des paiements à l'acte pour les visites de nuit, les vaccinations, les consultations en urgence pour des patients de confrères, la petite chirurgie, ainsi que des honoraires spécifiques plus globaux pour le suivi de la contraception, la surveillance des jeunes enfants, la prise en charge des grossesses.
- Des paiements aux résultats (« *target payments* »), pour le dépistage du cancer du col utérin (paiement lié au taux de femmes ayant eu un frottis dans les cinq années précédentes), pour la couverture vaccinale des enfants de deux ans et celle des enfants d'âge pré-scolaire.

La part de la rémunération qui est indépendante du niveau de service représente 65 % (50 % pour la capitation et 15 % pour les allocations de base), les paiements à l'acte 25 % et les paiements à la performance 10 %. En outre, toujours au Royaume-Uni, les médecins travaillant dans les régions rurales mal desservies peuvent dispenser les médicaments à leurs patients; ils reçoivent alors des honoraires spécifiques.

Les paiements incitatifs ont-ils eu les résultats escomptés? Il semble effectivement que les activités rémunérées à l'acte aient augmenté la plupart du temps, mais pas toujours (par exemple la proportion de médecins acceptant des suivis de grossesses complets a diminué de 1960 à 1980). Par ailleurs, on ne peut faire la part, dans cette augmentation d'activité, du rôle de l'incitation financière et d'autres facteurs liés à l'évolution générale de la pratique professionnelle et de la demande des patients (Whynes et Baines 1998).

Si, dans le National Health Service (NHS), des incitations à la productivité sont venues s'ajouter aux paiements à l'origine forfaitaires, on assiste aujourd'hui à l'introduction de forfaits dans des systèmes de paiement à l'acte, par exemple :

- Au Québec, pour les spécialistes en milieu hospitalier, la rémunération mixte a consisté à introduire une part forfaitaire et, en contrepartie, à ne payer les actes qu'à une fraction de leur tarif et à supprimer la facturation d'actes annexes.
- En France, pour être « médecins référents » les généralistes volontaires se voient attribuer un forfait annuel de 46 Euros pour le suivi général du patient, en plus de la tarification à l'acte (mais cette option n'a été choisie que par 10 % des généralistes et 1 % des patients).

Les formes les plus innovatrices de paiement mixte intègrent des éléments d'intéressement aux résultats (paiement à la performance). Ces résultats peuvent être sanitaires (comme dans l'exemple britannique ci-dessus), mais aussi financiers, c'est-à-dire liés à la maîtrise des coûts. Ces incitations sont utilisées par les organisations de *managed care* aux États-Unis en vue de réduire les adressages aux spécialistes et les recours à l'hospitalisation. Elles sont en général fondées sur un intéressement du médecin à la performance financière de l'organisation. Quel bilan peut-on tirer des expériences d'intéressement aux résultats?

L'intéressement aux résultats financiers

La réforme de 1991 au Royaume-Uni (*fundholding*) a introduit un intéressement financier pour les généralistes britanniques en ce qui a trait aux dépenses de médicaments et de chirurgie programmée des patients inscrits sur leur liste : le généraliste volontaire recevait un budget en début de période pour tous ses patients, et pouvait conserver l'éventuel surplus en fin d'année. Théoriquement, les surplus ne sont pas un revenu, car le généraliste doit les réinvestir dans sa pratique; mais, en réalité, il est difficile de séparer le revenu personnel du médecin de celui de sa pratique. L'évaluation de cette expérience a montré que les médecins ont effectivement réalisé des économies sur les budgets de médicaments, mais peu sur la chirurgie programmée, l'une des raisons étant la position de quasi-monopole des hôpitaux britanniques (Coulter 1992).

Une étude de Debrock et Arnould (1992) a mesuré l'effet des incitations financières axées sur le médecin, en contrôlant à l'aide des incitations existant au niveau du groupe. Leurs conclusions, pour les 35 organisations de *managed care* de l'Illinois examinées entre 1985 et 1987, sont que les incitations financières individuelles ont réduit le taux d'admission à l'hôpital de 16 % et diminué de moitié le nombre de consultations par bénéficiaire.

Si l'intéressement des médecins à limiter les dépenses semble efficace, le conflit d'intérêts qu'il risque de poser entre l'éthique professionnelle et l'intérêt financier du praticien est mal perçu par la profession et surtout par la population :

- Une enquête auprès de 766 médecins américains, dont 58 % faisaient l'objet d'incitations au niveau collectif ou individuel, a révélé que 57 % des praticiens ressentaient une pression pour limiter les adressages et 17 % pensaient que cela pouvait influencer

négalement sur la qualité de la prise en charge (Grumbach et coll. 1998, cités par Armour et coll. 2001). L'enquête donne aussi une idée des montants associés à ce genre de bonis : les médecins interrogés estimaient que les bonis représentaient environ 7 % de leur revenu médian.

- Bien que d'importance limitée, ces incitations financières inquiètent les patients, qui craignent qu'elles ne limitent l'autonomie professionnelle de leur médecin. Il y a peu d'indications empiriques sur l'étendue de cette préoccupation, si ce n'est les nombreux recours en justice portant sur les incitations financières offertes aux médecins. La législation américaine a d'ailleurs fixé en 1996 certaines limites aux contrats passés entre organisations de *managed care* et médecins et rendu leur publication obligatoire. Selon Hall, Dugan, Balkrishnan et coll. (2002), la publication des incitations accroît légèrement la confiance des patients si le médecin est payé par capitation, mais n'aurait aucun effet s'il est rémunéré à l'acte.

De nombreux auteurs, assureurs et HMO pensent aujourd'hui qu'il faut utiliser des incitations à la performance pour accroître la qualité, en valorisant des objectifs sanitaires plutôt que financiers. Selon Rosser et Kasperski (2002), les médecins de l'Ontario se déclarent convaincus par le passage à la capitation à condition que celle-ci soit mitigée par une rémunération à l'acte pour certains services, comme les gardes, et des primes pour l'atteinte d'objectifs de prévention.

L'intéressement aux résultats sanitaires

Ces formules ne portent pas directement sur des objectifs de santé – idée séduisante mais difficilement praticable – mais sur la prise en charge de la population suivie. La prévention, primaire (vaccination) ou secondaire (dépistage) en est le domaine d'application le plus fréquent.

L'exemple du Royaume-Uni cité ci-dessus semble montrer que ces « *target payments* » ont un impact positif : selon le ministère de la santé du Royaume-Uni (cité par Scott et Hall 1995), la proportion de généralistes atteignant l'objectif fixé est passée en cinq mois, pour la vaccination, de 55 % à 70 % (objectif : taux de vaccination de 90 % ou de 70 %) et de 53 % à 67 % pour les frottis (objectif : 80 % des femmes de 25 à 64 ans ayant eu un frottis dans les cinq ans). Certes, cette simple observation de résultats « avant-après » ne permet pas de neutraliser les autres facteurs explicatifs possibles et d'affirmer de manière rigoureuse que l'augmentation est due à l'incitation financière. C'est ce que souligne Armour et coll. 2001 à propos d'une autre expérience, tentée en Irlande du Nord, pour la vaccination des enfants : là encore, bien que le taux de vaccination ait fortement augmenté sur cette période, l'effet pur de cette mesure n'est pas dégagé d'autres effets concomitants.

Une expérience de bonis pour améliorer le dépistage des cancers, mise en place par une organisation de *managed care* pour des bénéficiaires du programme Medicaid, a été évaluée (Hillman et coll. 1998). Un boni semestriel allant de 10 % à 20 % de la capitation pour l'ensemble des femmes bénéficiaires était versé aux six cabinets qui réalisaient le plus fort taux de dépistage. Cinquante-deux cabinets ont été répartis aléatoirement entre un groupe

d'intervention (avec incitation) et un groupe témoin. L'étude n'a pas révélé de différence significative dans les taux de dépistage entre les deux groupes.

À l'inverse, des conclusions positives sont ressorties d'une expérience d'incitation à la vaccination grippale des personnes âgées de plus de 65 ans portant sur cinquante-quatre cabinets new-yorkais et évaluée par un essai contrôlé (Kouides et coll. 1998). En plus de la rémunération pour le vaccin (8,00\$), les médecins recevaient 0,80\$ par vaccin au-delà d'un taux de vaccination de 70 % pour les 65 ans et plus et de 1,60\$ pour un taux excédant 80 %. Le taux de vaccination dans le groupe avec boni (69 %) était supérieur de 6 points de pourcentage à celui du groupe témoin, et 52 % des cabinets recevant le boni avaient atteint un taux de vaccination de 70 %, contre 44 % dans le groupe témoin. Une analyse neutralisant un certain nombre de facteurs environnementaux permet aux auteurs de conclure que l'introduction des bonis a permis d'augmenter de 7 % le taux de vaccination des personnes âgées.

Il semble qu'une tendance récente consiste à lier le boni à plusieurs indicateurs de qualité. Dans une expérience menée par Aetna (Hanchak 1997), les groupes de médecins contractants ont reçu des bonis en fonction de leur score sur un indice de qualité des soins intégrant plusieurs dimensions, dont la satisfaction du patient, l'adéquation des traitements et l'efficacité. Les auteurs concluent à un effet positif sur toutes les dimensions de la qualité (baisse de 2 % du taux de césarienne, baisse de 25 % de la durée moyenne des séjours à l'hôpital, augmentation de 85 % des biopsies et taux de certification approchant 95 % à la fin du programme), à l'exception de la satisfaction du patient qui ne semble pas avoir augmenté.

En conclusion d'une revue de la documentation, Armour et coll. (2001) notent que les résultats des études sur l'effet des incitations sont contradictoires et permettent difficilement de porter un jugement définitif. Par contre, ils notent que les études traitent rarement à la fois d'effets de quantité et d'effets de qualité et ils y voient la reconnaissance d'une forme d'arbitrage entre la maîtrise des coûts et la qualité de soins.

Combinaisons d'incitations différentes au niveau individuel et collectif

Une deuxième approche consiste à tenter, non pas de contrebalancer différents incitatifs au sein de la rémunération individuelle, mais de compenser les incitations individuelles par des incitations collectives.

Elle s'illustre par exemple dans les systèmes d'enveloppe limitative globale associée au paiement à l'acte. L'Allemagne a opté pour un tel système dès le début des années 80 et d'autres pays ou régions ont adapté ce principe, selon des modalités variables en fonction de leur contexte propre.

Les médecins allemands sont rémunérés à l'acte mais soumis au principe d'une modération collective : les paiements à l'acte de l'ensemble des médecins d'un Land sont plafonnés, et le montant est réparti entre les praticiens par les syndicats de médecins, au prorata de l'activité de

chacun, exprimée en points sur une nomenclature d'actes. En d'autres termes, la valeur unitaire du point n'est pas connue à l'avance mais est déduite *a posteriori* en divisant l'enveloppe totale par le nombre cumulé de points de tous les médecins (Busse 2000).

La rémunération directe des praticiens par les caisses, sous forme de paiement cumulé sur une période donnée (trimestre en Allemagne) conditionne le caractère opératoire d'un tel système. Cependant, des systèmes où ces conditions ne sont pas remplies ont aussi instauré des plafonds de dépenses (Québec, France, États-Unis pour Medicare) : la différence est que l'ajustement sur le tarif unitaire ne peut se faire automatiquement en temps réel, mais doit être décidé – et négocié – pour l'année suivante, en fonction des résultats constatés. Cette modalité a évidemment toutes les chances d'être moins efficace sur le plan du strict respect de l'enveloppe financière.

En théorie, ce type de mécanisme suscite un conflit, pour chaque médecin, entre l'intérêt collectif à une modération quantitative (une valeur du point plus élevée et un niveau de revenu globalement inchangé pour quantité de travail moindre) et l'intérêt individuel qui reste de maximiser son volume d'activité propre à un prix unitaire donné. Il favorise donc les comportements de type « passager clandestin », chaque médecin se disant qu'il pourra profiter à titre personnel de la modération collective s'il augmente son volume d'activité, ce qui au total conduit à une inflation des volumes contrairement à l'objectif recherché.

Cela est lié, évidemment, à la dilution de la responsabilité au sein d'un collectif très important. Individuellement, chaque médecin n'a aucune chance d'influer sur la valeur unitaire par un comportement vertueux, mais il a, au contraire, tout intérêt à être ce « passager clandestin » du grand collectif anonyme. Nous reviendrons sur la question de la taille des groupes de médecins ci-dessous.

Les attendus théoriques de ce genre de système – inflation continue des volumes et baisse corrélative des prix unitaires – semblent bien confirmés par l'observation en Allemagne (encore que l'on puisse débattre, s'agissant de l'augmentation des volumes, de la part liée à ces comportements opportunistes des médecins ou à d'autres facteurs – demande, progrès technique, etc.).

Il faut donc un mécanisme complémentaire pour contrebalancer cette incitation à des comportements individuels trop « productivistes ». Celui-ci peut relever :

- Du contrôle : Il est présent en Allemagne, où l'activité individuelle est examinée par les médecins des caisses. Au passage, on peut penser que ce contrôle sera d'autant plus efficace qu'il est organisé par la profession elle-même; la faisabilité de ces dispositifs dépend de la disposition du corps médical à co-gérer la contrainte financière avec les régulateurs, comme le montre *a contrario* l'exemple français.
- De l'incitation, comme au Québec, avec des tarifs dégressifs au-delà d'un seuil quantitatif d'actes, ce qui écrête les très fortes activités.

Le groupe de médecins comme lieu d'intégration des objectifs cliniques et économiques

Dans différents pays, on assiste, semble-t-il, à une évolution convergeant vers un exercice médical plus collectif, sous des diverses formes – réseaux, groupes monodisciplinaires ou groupes pluridisciplinaires.

Cette évolution émerge, depuis le milieu des années 90, dans des contextes nationaux aussi contrastés que le NHS anglais et le *managed care* américain (Bocognano et coll. 1998). Au Royaume-Uni, la réforme de 1997, en instituant les *Primary Care Groups* ou *Trusts* (PCG/T), a généralisé les premiers regroupements spontanés de généralistes *fundholders*; dans les contextes locaux les plus concurrentiels aux États-Unis, des groupes de médecins se sont développés très rapidement, pour passer des contrats avec les assureurs en soins de santé cherchant à constituer des réseaux, voire dans les cas extrêmes intégrer eux-mêmes les fonctions de l'assureur.

Dans les deux cas, la perspective est de faire de ces groupes des réseaux très structurés, possédant leur propre infrastructure de gestion.

Ces groupes se voient parfois confier une responsabilité financière et une délégation de gestion. Ainsi les PCG/T britanniques sont-ils rémunérés par des paiements globaux, à charge pour eux de répartir le financement entre les différents cabinets médicaux qui les composent. Prenant la suite de l'expérience des *fundholders*, les budgets alloués englobent non seulement la rémunération des médecins, mais aussi les prescriptions et les hospitalisations.

Les PCG/T sont donc à la fois responsables du respect d'un budget global, des arbitrages entre différentes composantes de ce budget, et de la répartition des honoraires et budgets de prescription entre leurs membres.

Pour cela, ils doivent développer une expertise leur permettant d'optimiser l'utilisation des ressources. Ainsi, dans le domaine de la prescription médicamenteuse, tous les PCG/T surveillent les prescriptions de leurs membres à l'aide des tableaux indicatifs fournis par le palier national, presque tous ont un conseiller pharmaceutique, 70 % utilisent des référentiels communs de prescription et 49 % ont produit leur propre formulaire (Wilkin, Gillam et Coleman, 2000/2001).

Cet exemple illustre l'intérêt des groupes de médecins, qui réside dans la possibilité de lier la responsabilité financière et la gouvernance clinique dans un collectif :

- Géré par les professionnels, et ayant donc à la fois la compétence technique pour faire les meilleurs choix et une éthique de la qualité.
- De taille suffisante pour se doter d'un certain nombre d'outils d'expertise et d'évaluation et, en même temps, de taille suffisamment restreinte pour que chaque professionnel se sente impliqué.

L'engouement pour les groupes médicaux, jugés les mieux aptes à mettre en œuvre une gestion efficace des processus de soins, et dès lors de concilier qualité des soins et utilisation

économique des moyens, a également été très fort dans le système de *managed care* américain. Dans les zones les plus concurrentielles, on a vu durant la deuxième moitié des années 90 un développement très rapide de ces groupes médicaux et de leur capacité de gestion et de négociation, tant avec les assureurs en amont qu'avec les hôpitaux en aval.

Au sein de ces réseaux, des groupes par spécialité peuvent se voir attribuer des budgets globaux dont ils gèrent la répartition entre les praticiens individuels. Ceux-ci peuvent être payés à l'acte, mais le groupe met en place des procédures (revue par les pairs, vérification de l'utilité des soins, etc.) qui sont plus opérationnelles au sein d'une même spécialité que dans des groupes multidisciplinaires (Robinson 2001b). Dans ce cadre, peuvent être mis en œuvre des mécanismes de capitation, dits « capitation de contact », assez novateurs pour des spécialistes : dès qu'un patient est confié par un généraliste à un médecin du groupe, celui-ci reçoit une capitation pour une durée de 12 mois, durée pendant laquelle il prend en charge le patient comme il l'entend.

Ce principe de rémunération est plus adapté aux spécialités dont l'activité est à peu près homogène (cardiologie, gastro-entérologie, etc.). Il peut être rajusté pour tenir compte de la lourdeur des cas, par exemple en pondérant les effectifs de patients pour certains diagnostics ou certains actes, ou en prévoyant la rémunération à l'acte pour certaines procédures (Franck 1999). Ce mode de paiement a été utilisé dans plusieurs HMO et les spécialistes du *managed care* lui prédisaient récemment un bel avenir (Carlson 1997). L'expérience suédoise semble confirmer que l'incitation collective peut avoir un profond impact, même si le revenu individuel n'est pas touché par le changement de mode de paiement : certains hôpitaux sont passés à un mode de paiement à la pathologie, alors que d'autres conservaient l'enveloppe annuelle fixe; mais, dans les deux cas, les médecins étaient salariés de l'hôpital. Il en résulte que les médecins des hôpitaux rémunérés à la pathologie ont notablement réduit les durées de séjour, même si le coût total n'a pas diminué dans ces hôpitaux (Forsberg, Axelsson et Arnetz 2001).

Les expériences décrites ci-dessus concernent des réseaux très formalisés ayant de fortes délégations de responsabilité. Mais l'idée de promouvoir un exercice plus collectif de la médecine, sous forme de réseaux ou de groupes, se développe également dans d'autres systèmes, de manière moins structurée (France, Allemagne, certaines provinces canadiennes avec les expériences de groupes de médecine familiale). Dans ce cas, l'objectif est plutôt d'améliorer la coordination et la continuité des soins, qu'une pratique individuelle et rémunérée à l'acte ne favorise pas spontanément. Mais l'idée sous-jacente est la même : entre le médecin isolé et la profession dans son ensemble, le groupe de médecins, localement implanté, partageant des compétences et une culture commune, peut être un levier de changement des pratiques dans le sens recherché.

Cependant, l'évolution récente aux États-Unis donne à réfléchir sur les excès de cette tendance. La Californie, « royaume du groupe médical », a connu ces dernières années plusieurs faillites et voit aujourd'hui les groupes se défaire et la contractualisation individuelle reprendre de l'ampleur. Le recul des groupes va de pair avec le recul de la capitation comme mode de paiement (Managed Care Outlook 2001) : le pourcentage des groupes ayant au moins un contrat à la capitation, qui était en hausse entre 1998 et 1999 (66 % à 72 %) a fléchi en 2000 (55 %).

Les facteurs qui expliquent la fragilité des groupes et leur absence de pérennité (déséconomies d'échelle, difficulté à intégrer des logiques d'assurance, demande de libre choix des patients, etc.) montrent bien que la conciliation des objectifs de maîtrise des coûts et de qualité par la vertu de la « gouvernance clinique » collective a ses limites et qu'une responsabilisation financière extrême n'est peut-être pas le moyen le plus propice de tirer le meilleur parti de ces organisations collectives qui paraissent malgré tout, dans beaucoup de systèmes, constituer une voie de progrès.

Incitations non financières

Si les groupes de médecins peuvent harmoniser la maîtrise des coûts et la qualité des soins, ce n'est pas seulement par l'alchimie des incitations financières collectives et individuelles auxquelles ils sont confrontés : c'est bien parce qu'ils associent aux incitations financières d'autres mécanismes incitatifs.

Ceux-ci sont de plusieurs ordres. Le premier instrument est l'**élaboration et la diffusion de protocoles cliniques et de recommandations**, aujourd'hui communes à la plupart des systèmes publics ou privés. Ils peuvent être nationaux ou locaux, avoir un caractère indicatif ou obligatoire (formulaires pour les médicaments, par exemple). L'efficacité comparée des différentes modalités de diffusion des références (formation, entretiens, rappels informatiques) sur leur application par les médecins a fait l'objet d'une abondante documentation, suite aux recommandations de la *Cochrane Collaboration on Effective Professional Practice*. Il serait trop long de restituer et de commenter cette documentation dans le cadre du présent rapport, mais nous pouvons en retenir les enseignements principaux, tels que résumés par Durieux (1998) : aucune étude ne décèle d'impact significatif de la diffusion de normes par matériel imprimé ou par formation médicale continue; en revanche, la diffusion par des leaders d'opinion, ou, mieux encore, par des visites confraternelles a un impact significatif sur les pratiques des médecins.

En congruence avec ce premier résultat, **l'évaluation de la pratique individuelle**, notamment au regard de ces référentiels, ou par rapport aux confrères (restitution d'information au médecin, revue par les pairs, etc.) est un levier de modification des comportements individuels.

Le réseau de soins ou le groupe de médecins constitue un niveau privilégié d'application de ces incitations, pour les raisons invoquées ci-dessus. Ce sera d'autant plus le cas que ces réseaux sélectionnent leurs membres en fonction de leur adhésion aux valeurs du groupe – sur les plans clinique, économique et culturel. Un pas supplémentaire peut être franchi dans les systèmes à assureurs multiples avec le **conventionnement sélectif** (capacité pour le payeur de ne prendre en charge qu'un nombre limité de professionnels), qui constitue évidemment une puissante incitation non financière.

Conclusion : Quels enseignements pour le contexte canadien?

Le Canada occupait la cinquième place des pays de l'OCDE en 1998 pour la part du PIB consacrée aux dépenses de santé, et le rapport intermédiaire de la Commission sur l'avenir des soins de santé au Canada note que le problème réside plutôt dans l'instabilité de l'évolution de ces dépenses que dans leur niveau absolu (OCDE 2001, Commission 2002).

Toujours en 1998, le Canada se classait au septième rang pour les dépenses en soins ambulatoires par personne, avec un niveau (333\$PPA) représentant le tiers de celui des États-Unis, 80% de celui de l'Allemagne, mais supérieur de 40 % à celui de la France. L'évolution des prix des services de médecins au Canada sur la décennie 1985-1996 coïncide avec celle d'autres pays de l'OCDE (à l'exclusion des États-Unis) et la progression des volumes, assez forte au début par comparaison avec les autres pays, semble avoir été freinée à partir de 1995.

Le Canada ne semble donc pas se distinguer par un problème spécifique de maîtrise des coûts des soins de santé⁶. En revanche, le rapport intermédiaire de la Commission souligne la grande insatisfaction des Canadiens au regard de l'accessibilité et de la qualité du système de soins.

Pour résoudre ces problèmes, faut-il remettre en cause la rémunération à l'acte? On a vu que le paiement par capitation peut faire peser un risque sur la qualité des soins et l'accessibilité, d'autant plus que le médecin ne redoute pas la concurrence, par exemple en cas de pénurie de médecins. Par contre, en modulant la rémunération du médecin selon le type de population qu'il dessert, la capitation peut permettre de traiter des problèmes particuliers comme la répartition géographique inégale des médecins sur le territoire.

Le débat actuel sur l'avenir du système de santé canadien est sans doute l'occasion d'un débat sur le mode de paiement. D'autant que d'après le dernier sondage réalisé par l'Association médicale canadienne, si 62 % d'entre eux tirent la quasi-totalité de leurs revenus de paiements à l'acte, seulement 37 % préfèrent ce mode de rémunération aux autres modes (Martin 2000).

Questions spécifiques

Quel est l'impact des mécanismes de paiement des médecins aussi bien sur l'utilisation des services médicaux que sur l'utilisation générale et les coûts globaux du système de santé?

Comment le paiement à l'acte des médecins influence-t-il l'utilisation et les coûts du système par rapport à d'autres mécanismes de paiement, et quelles sont les différences que l'on peut observer à ce chapitre?

Il n'existe pas de réelle alternative aux systèmes de paiements canoniques que sont le paiement à l'acte, la capitation et le salariat. Les innovations récentes en sont plus des adaptations à la marge. On synthétise l'état des connaissances sur l'impact des modes de paiement sur les coûts des soins, selon leur degré de forfaitisation : le paiement à l'acte permet au médecin d'augmenter le volume des services médicaux qu'il fournit au patient, et limite donc la portée des modes de régulation par les prix unitaires.

À l'inverse, les paiements forfaitisés (capitation et salariat) limitent les volumes, mais peuvent peser sur la qualité des soins rendus, limiter l'accès aux soins pour les populations vulnérables et avoir des conséquences imprévues sur le niveau des dépenses. L'ampleur de l'impact sur la qualité et l'accès dépend de manière cruciale de la capacité du patient d'observer l'effort du médecin, capacité sur laquelle il est difficile de se prononcer empiriquement. L'effet imprévu sur les coûts passe par la prescription (la forfaitisation incite le médecin de première ligne à adresser et prescrire plus) et par les effets à long terme (si le médecin réduit le service à court terme, le système peut en payer les conséquences à long terme).

Les réponses concrètes apportées visent à combiner les mécanismes, de façon à ce que les défauts se compensent. Les paiements en fonction de résultats sanitaires (objectifs intermédiaires) permettent notamment d'améliorer la capacité du mode de paiement à la capitation de tenir compte de la diversité du travail du médecin, et de combiner des objectifs de santé publique, d'une part, et d'accès et de qualité, de l'autre. Les délégations de responsabilité à la profession médicale (le régulateur fixe une enveloppe, gérée ensuite de manière décentralisée) se heurtent à des comportements opportunistes si la décentralisation est trop faible, et à des conflits d'intérêts si elle est trop poussée (groupes de médecins trop restreints). Après avoir été présentée comme la solution de l'avenir, la délégation à des groupes de médecins semble aujourd'hui marquer le pas, au profit des incitations non financières, centrées sur l'information du patient et du médecin.

*De quelles données dispose-t-on sur l'impact des changements en matière de mécanismes de paiement ou de mesures incitatives sur l'utilisation et les coûts?
Où et comment se produisent les changements les plus marqués?*

Les données quantitatives sont rares, opportunistes (on utilise des expériences faites pour d'autres buts) et toujours affectées de biais. Les changements exogènes notables se sont produits au Royaume-Uni, avec l'introduction du *fundholding* et des PCT, au Danemark (expérience de Copenhague) et en Allemagne (enveloppe et point flottant). Aux États-Unis, il y a eu de

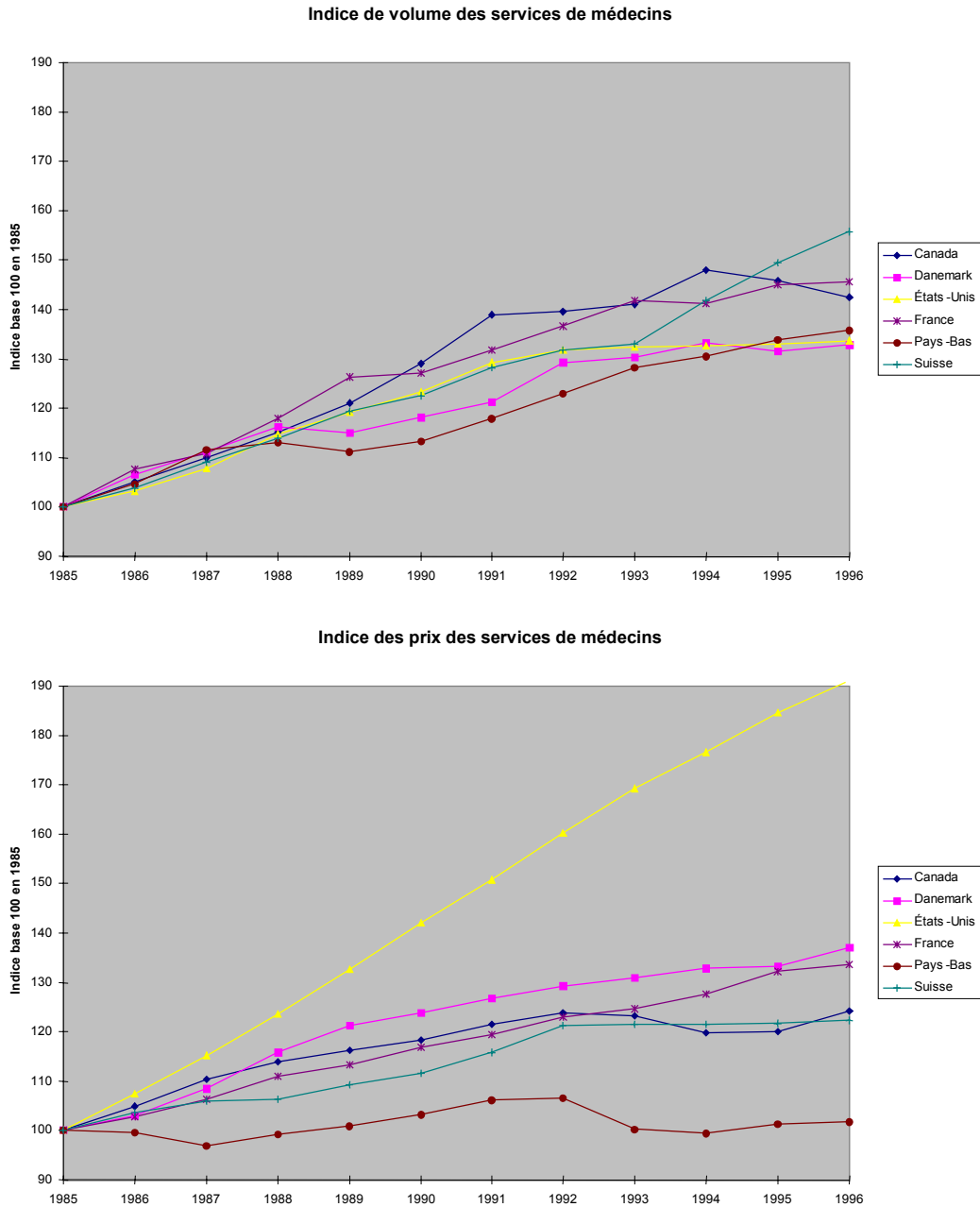
nombreuses expériences sur des échantillons réduits (une clinique ou un régime), qui sont difficilement généralisables. La réforme des tarifs de Medicare a aussi constitué un terrain d'expérimentation, mais la difficulté principale est qu'on mêle niveau et mode de rémunération.

Il ressort des données quantitatives empiriques que le passage de la capitation au paiement à l'acte augmente les volumes et diminue la prescription, sans qu'on ait de preuve réelle de l'effet réciproque. Le bilan sur la qualité est plus ambigu, même si la satisfaction des patients semble défavorable au régime forfaitaire. Les expériences de paiement au résultat semblent modérément concluantes quant à l'impact des incitations sur le résultat recherché. Outre les évaluations quantitatives, on trouve des bilans qualitatifs, basés sur les opinions des acteurs; le suivi conjoncturel de l'histoire des groupes de médecins aux États-Unis et de leurs arrangements internes fournit de nombreux résultats, qui sont toutefois difficilement généralisables à des contextes où il n'y a qu'un seul payeur.

Quelles sont les mesures incitatives non monétaires qui peuvent affecter l'utilisation et les coûts du système? Comment fonctionnent-elles?

Deux innovations principales sont proposées : aider les médecins en leur fournissant une information validée et synthétisée sur les protocoles cliniques et l'efficacité, notamment en remplaçant la promotion provenant de l'industrie des biens médicaux et aider les médecins et les patients à se faire une idée de la variation des pratiques. Ce deuxième genre d'incitation permet de renforcer le contrôle des patients sur l'effort du médecin, ce qui peut augmenter la variabilité (adaptation des pratiques à la demande) ou la diminuer (standardisation). Mais il ne s'agit pas seulement de nommer et de dénoncer; les médecins peuvent aussi trouver avantage à comparer leur pratique à celles de leurs confrères. L'impact de ces incitations non monétaires dépend fortement du contexte, notamment du degré d'organisation de la profession et de la capacité des médecins de diffuser la conformité aux normes de bonnes pratiques parmi leurs pairs.

Figure 1
Évolution des prix et des volumes* des services de médecins, 1985-1996



* Les volumes correspondent aux dépenses en soins de médecins ambulatoires, déflatées par l'indice de prix correspondant.
Source : OCDE, 2001.

Notes

- 1 Augmenter la quantité peut signifier modifier la structure des services offerts, en augmentant la part des actes mieux rémunérés. Rochaix (1993) montre, en prenant l'exemple québécois entre 1977 et 1983, que les gels d'honoraires ont conduit à une augmentation significative de la part des consultations dites « lourdes ».
- 2 En régime salarial, le médecin peut éventuellement être sanctionné par son employeur. Dans le cas de la capitation, il sera sanctionné par le patient si la sous-qualité est perçue et si les professionnels sont en situation de concurrence pour attirer la clientèle.
- 3 L'adressage est l'action par laquelle un professionnel dirige son patient vers un autre professionnel.
- 4 Ces données permettent de contrôler les biais dans les études par comparaisons en captant les effets propres à chaque individu; en revanche, parce qu'il s'agit de données complexes à obtenir, les échantillons sont moins souvent représentatifs.
- 5 Une autre étude sur le même ensemble de données montrerait un impact significatif et important du paiement à l'acte sur le volume d'examens d'urine (Kristiansen et Hjortdahl 1992). Mais Scott et Hall (1995) soulignent que cet effet est mesuré sans corriger le biais lié à la non indépendance des observations (on observe plusieurs contacts par patient).
- 6 Ce diagnostic repose sur les données disponibles au moment de la rédaction de cet article (c'est-à-dire 1996 ou 1998 selon les séries). Il est possible que les experts canadiens portent un autre jugement à la lumière de données plus récentes sur l'évolution des dépenses, des prix et des volumes produits.

Bibliographie

- Armour, Brian S., et coll. 2001. The effect of explicit financial incentives on physician behaviour, *Archives of Internal Medicine*, 161(10), 1261-1266.
- Bocognano Agnès, et coll. 1998. *Mise en concurrence des assurances dans le domaine de la santé : théorie et bilan des expériences étrangères*, CREDES-INSEE.
- Brunelle, Yvon, et Howard Bergman. 1999. Capitation, efficacité et rationnement, *Ruptures, revue transdisciplinaire en santé*, 6(2), 270-288.
- Busse, Rheinarth. 2000. *Health Systems in Transition – Germany*, European Observatory on Health Care Systems.
- Carlson, Bob. 1997. “Capitation: State of the art”, <<http://www.managedcaremag.com/archives>>.
- Commission sur l'avenir des soins de santé au Canada. 2002. *Préparer l'avenir des soins de santé*, Rapport d'étape, février.
- Coulter, Angela. 1992. The interface between primary and secondary care, dans *Hospital Referrals*, sous la direction de M. Roland et A. Coulter, Oxford General Practices Series 22, Oxford University Press.
- Davidson, Stephen, M. Manheim, S, Werner et coll. 1992. Prepayment with office-based physicians in publicly funded programs: results from the children's Medicaid program, *Pediatrics*, 89, 761-767.
- Debrock, Larry, et Richard J. Arnould. 1992. Utilization control in HMOs, *Quarterly Review of Economic Finance*, 32, 31-53.
- Durieux, Pierre. 1998. Les actions de diffusion et de formation, dans *Comment améliorer les pratiques médicales? Approche comparée internationale*, sous la coordination scientifique de Pierre Durieux, Institut d'Études des Politiques de Santé, 11-18.
- Emery, J.C. Herbert, Chris Auld et Mingshan Lu. 1999. *Paying for physician services in Canada: The institutional, historical and policy context*, Institute of Health Economics, Working Paper 99-6.
- Escarce, José. 1996. Using physician practice cost functions in payment policy: the problem of endogeneity bias, *Inquiry*, 33, 66-78.
- Epp, Michael, et coll. 2000. The impact of direct and extra billing for medical services: Evidence from a natural experiment in British Columbia, *Social Science and Medicine*, 51, 691-702
- Evans, Robert, G. 1974. Supplier induced demand: Some empirical evidence and implications, dans *The Economics of Health and Medical Care* McMillan, sous la direction de M. Perlman 162-173.
- _____. 1998. New bottles, old wine: right and wrong in physician supply, *Canadian Medical Association Journal*, 158, 757-759.
- Ferrall, Christopher, Allan W. Gregory et William G. Tholl. 1998. Endogenous work hours and practice patterns of Canadian physicians, *Revue canadienne d'économique*, 31(1), 1-27.

- Forsberg, Ewa, Runo Axelsson et Arnetz Bengt, 2001, Financial incentives in health care. The impact of performance-based reimbursement, *Health Policy* 58(3): 243-262.
- Franck, Clifford R. 1999. Can the pie be divided more fairly?, *Contact capitation for specialists*, 5(2), <http://www.cost-quality.com/5,2art3.html>.
- Gerdtham, Ulf-G., Jes Søgaard, Fredrik Andersson et Bengt Jönsson. 1992. Econometric analysis of health expenditure: A cross-section study of the OECD countries *Journal of Health Economics*, 11, 63-84.
- Gerdtham, Ulf-G, Bengt Jönsson, M. MacFarlan et H. Oxley. 1998. The determinants of health expenditure in OECD countries, dans *Health, the Medical Profession, and Regulation*, sous la direction de Peter Zweifel, Kluwer.
- Gillett, James, Brian Hutchison et Stephen Birch. 2001. Capitation and primary care in Canada: Financial incentives and the evolution of health service organizations, *International Journal of Health Services*, 31(3), 583-603.
- Glied, Sherry. 1998. Payment Heterogeneity, Physician Practice, and Access to Care, *American Economic Review*, 88(2), 127-131.
- Gosden, Toby, Frode Forland, Ivar Sombo Kristiansen et coll. 2001. Impact of payment method on behaviour of primary care physicians: A systematic review, *Journal of Health Services Research and Policy*, 10(1), 44-55.
- Grumbach, Kevin, Dennis Osmond, Karen Vranizan et coll. 1998. Primary care physicians' experience of financial incentives in managed care systems, *New England Journal of Medicine*, 339, 1516-1521.
- Grytten, Jostein, et Rune Sorensen. 2001. Type of contract and supplier-induced demand for primary physicians in Norway, *Journal of Health Economics*, 20, 379-393.
- Hall, Mark, A., Elizabeth Dugan, Rajesh Balkrishnan et coll. 2002. How Disclosing HMO Physician Incentives Affects Trust, *Health Affairs*, 21(2), 19-27.
- Hanchak, N. A. 1997. A performance-based compensation model for obstetricians/gynecologists, *Clinical Obstetrics and Gynecology*, 40, 437-445.
- Hemenway, David. 1995. Financial incentives for childhood immunization, *Journal of Policy Analysis and Management*, 14, 133-139.
- Hickson, B., A. Altemeier, et M. Perrin. 1987. Physician reimbursement by salary of fee-for service: Effect on physician practice behaviour in a randomized prospective study, *Pediatrics*, 80, 344-350.
- Hillman, Alan, L., Kimberly Ripley, Neil Goldfard et coll. 1998. Physician financial incentive and feedback: Failure to increase cancer screening in Medicaid managed care, *American Journal of Public Health*, 88, 1699-1701.
- Hutchinson, James M., et Robert N. Foley. 1999. Method of physician remuneration and rates of antibiotic prescription, *Canadian Medical Association Journal*, 160, 1013-1017.

- Jelovac, Izabela. 2001. Physicians' payment contracts, treatment decisions and diagnosis accuracy, *Health Economics*, 10, 9-25.
- Josseran, L., G. King et J. Chaperon. 2000. Les primary care physicians américains, *Journal d'économie médicale*, 18(6), 375-387.
- Kouides, R.W., N.M. Bennett, B. Lewis, J.D. Cappuccio, WH. Barker et F.M. LaForce. 1998. Performance-based physician reimbursement and influenza immunization rates in the elderly, *American Journal of Preventive Medicine*, 14, 89-95.
- Kralewski John E. et coll. 2000. The Effects of Medical Group Practice and Physician Payment Methods on Costs of Care, *Health Services Research*, 35(3), 591-613.
- Krasnik, Allan P. Groenwegen et Poul-A. Pedersen. 1990. Changing remuneration system: Effects on activity in general practice, *British Medical Journal*, 300, 1698-1701.
- Kristiansen, Iwar S., et K. Holtedahl. 1993. The effect of the remuneration system on the general practitioner's choice between surgery consultations and home visits, *Journal of Epidemiology and Community Health*, 47, 481-484.
- Kristiansen, Iwar S., et Gail Mooney. 1993. The general practitioner's use of time: Is it influenced by the remuneration system?, *Social Science and Medicine*, 37, 393-399.
- Kristiansen, Iwar S., et Per Hjortdahl. 1992. The general practitioner and laboratory utilization: Why does it vary?, *Family Practice*, 9, 22-27.
- Managed Care Outlook. 2001. "Interest in capitation back off", <<http://www.managedcaremag.com/archives>>.
- Mc Guire, Thomas G. 2000. Physician agency dans *Handbook of Health Economics*, sous la direction de Anthony J. Culyer et Joseph P. Newhouse volume 1, p. 463-536.
- Martin, Shelley. 2000. CMA survey shows fee-for-service not dead yet, *Canadian Medical Association Journal*, 163(5), 601.
- Miller, Robert H., et Harold S. Luft. 1997. Does managed care lead to better or worse quality of care?, *Health Affairs*, 16(5), 7-25.
- Mulligan, Pamela. 2002. La capitation : une direction mal avisée dans la réforme des soins de première ligne, *Le Médecin de famille canadien*, 48, 244-246.
- Namovicz-Peat, S., J. Swann et K. Tureski. 2001. *HMO & PBM Strategies for Pharmacy Benefits: Recent Results, Current Practices, Future Plans*, AIS - Atlantic Information Services.
- OCDE, 2001. *Eco-santé OCDE 2001*, CREDES-OCDE.
- Philips, Karen, et William R. Swan. 1996. *Health care system in transition: Canada*, version préliminaire, Organisation mondiale de la santé – Bureau régional pour l'Europe, Copenhague.

- Pricewaterhouse Coopers LLP. 2001. *Study on Pharmaceutical Benefit Management*, disponible sur le site de Health Care Financing Administration.
- Redbook 2002. <http://www.nhs.uk/redbook/2.htm>.
- Rice, Thomas. 1993a. A model is only as good as its assumptions: A reply to Peele. *Journal of Health Economics*, 12, 209-211.
- _____. 1993b. Demand curves, economists, and desert islands: A response to Feldman and Dowd. *Journal of Health Economics*, 12, 201-204.
- Robinson, James C. 2001a. Theory and practice in the design of physician payment incentives, *The Milbank Quarterly*, 79(2), 148-177.
- Robinson, James C. 2001b. Physician organization in California: Crisis and opportunity, *Health Affairs*, 20(4), 81-96.
- Rochaix, Lise. 1993. Financial Incentives for Physicians: the Quebec Experience, *Health Economics*, 2, 163-176.
- Rosser, Walter, et Jan Kasperski. 2002. En faveur d'une rémunération mixte, *Le Médecin de famille canadien*, 48, 247-249.
- Sandier, Simone. 1990. Le paiement des médecins dans quelques pays de l'OCDE, dans *Les systèmes de santé, A la recherche d'efficacité*, Paris, OCDE, 1990.
- Santé Canada. 2001. Loi canadienne sur la santé, *Rapport annuel 2000-2001*.
- Scott, Anthony, et Jane Hall. 1995. Evaluating the effects of GP remuneration: problems and prospects, *Health Policy*, 31, 183-195.
- Scott, Anthony. 2000. Economics of general practice, dans *Handbook of Health Economics*, sous la direction de Anthony J. Culyer et Joseph P. Newhouse volume 1, p.1175-1200.
- Showstack, Jonathan A., Deborah W. Garnick, Kenneth E. Rosenfeld et coll. 1987. Episode-of-care physician payment: A study of coronary artery bypass graft surgery, *Inquiry*, 24, 376-383.
- Stearns, Sally C., Barbara L. Wolfe et David A. Kindig. 1992. Physician responses to fee-for-service and capitation payment, *Inquiry*, 29, 416-429.
- Tai-Seale, Ming, Thomas H. Rice and Sally C. Stearns. 1998. Volume responses to Medicare payment reductions with multiple payers: A test of the McGuire-Pauly model, *Health Economics*, 7, 199-219.
- Whynes, David, et Darrin L. Baines. 1998. Income-based incentives in UK general practice, *Health Policy*, 43, 15-31.
- Wilkin, David, Steve Gillam et Anna Coleman. 2000/2001. *The national tracker survey of primary care groups and trusts*, www.kingsfund.org.uk.