

Rémunération des médecins de ville

Ce cours vous est proposé par Jérôme Wittwer, Professeur des universités en économie, Université de Bordeaux, et AUNEGe, l'Université Numérique en Économie Gestion.

Table des matières

Préambule	2
Comment sont déterminées les rémunérations des médecins ? Mécanismes concurrentiels ou prix administrés ?	2
Rémunérations des médecins : quels enjeux pour le financeur ?	3
Les différentes modalités de rémunérations des médecins et leurs propriétés	6
Le paiement à l'acte	7
La capitation (paiement par tête)	8
Discussion sur la tarification optimale	11
Références	15

Préambule

Objectifs d'apprentissage : assimiler les caractéristiques des principales modalités de rémunération des médecins. Savoir analyser les propriétés incitatives des différentes modalités de rémunérations des médecins.

Dans l'ensemble de ce chapitre nous nous intéressons à la rémunération des médecins (généralistes ou spécialistes) exerçant dans le secteur ambulatoire communément appelé médecine de ville. Par commodité, nous parlerons indifféremment de médecins ou de professionnels de santé l'ensemble des développements de ce chapitre pouvant s'appliquer à la rémunération de tout professionnel de santé intervenant dans le secteur ambulatoire.

Comment sont déterminées les rémunérations des médecins ? Mécanismes concurrentiels ou prix administrés ?

La rémunération des médecins ne répond que marginalement à une logique marchande simple qui verrait les prix des actes des médecins se fixer sur un marché des soins lieu de rencontre des patients et des médecins. Ceci est lié au fait que dans la plupart des pays les soins sont couverts, dans des proportions variables, par des assurances santé qui rémunèrent directement les médecins ou remboursent les patients quand ceux-ci avancent leurs frais de santé (comme c'est le cas en France). La question de la rémunération des médecins est ainsi très étroitement liée à celle de l'organisation du financement des soins.

Elle dépend plus précisément de l'intensité de la concurrence entre assureurs et de leur capacité à contractualiser avec les offreurs de soins. Dans une véritable organisation concurrentielle du financement des soins les assureurs sont en concurrence sur le marché de l'offre de soins dans le sens où ils contractualisent avec les offreurs de soins sur les prix des actes et les modalités de rémunération. Le prix pratiqué par un médecin pour un patient dépend ainsi de l'assureur par lequel il est couvert. En quelque sorte les assureurs se substituent aux patients pour mettre les offreurs de soins en concurrence. C'est la logique qui préside par exemple aux Etats-Unis. Dans une telle organisation les rémunérations des médecins (montant et modalités) sont ainsi le résultat des mécanismes concurrentiels sur le marché sur lequel interagissent assureurs et offreurs de soins. L'efficacité de ces modes de régulation des prix des soins dépend de l'efficacité des mécanismes concurrentiels et en particulier du pouvoir de marché des offreurs de soins c'est-à-dire de leur capacité à négocier des prix élevés avec les assureurs.

Si cette logique concurrentielle est à l'œuvre dans de nombreux pays son influence est dans les faits souvent périphérique. En effet, les prix des soins sont généralement administrés. Ils peuvent être administrés soit directement par l'état (ou des entités administratives autonomes) comme dans les systèmes beveridgiens (financement des soins par l'Etat et donc l'impôt), soit par des caisses publiques (alimentés par des cotisations sociales) après négociations avec les professionnels de santé, comme dans les systèmes bismarckiens. Des prix administrés signifient que les professionnels de santé sont soumis à des modalités de paiement auxquelles ils ne peuvent déroger. Cependant dans la plupart des pays dans lesquels les prix sont administrés une part de l'activité de soins y échappent. Dans le cas français, certains médecins ont ainsi la possibilité de pratiquer des dépassements d'honoraires c'est-à-dire de pratiquer des prix plus élevés que les prix administrés (aussi appelé tarifs). La couverture des dépassements peut être assurée par des assurances complémentaires privées qui n'ont que marginalement une action de contractualisation sur les prix avec les professionnels de santé (voir chapitre ???). Dans d'autres pays, comme en Allemagne par exemple, une part de l'activité de soins échappe complètement à l'administration des prix. Les patients recourant à ces soins doivent alors être couverts par des assurances dites privées qui sont en charge de la négociation des prix avec les offreurs de soins.

On voit ainsi que logique concurrentielle et celle de l'administration des prix se superposent ou se juxtaposent dans les systèmes de santé. La rémunération des médecins est le résultat de logiques différentes en fonction du type ou du secteur d'activité des médecins.

Les objectifs du financeur sont bien sûr différentes qu'il soit une assurance privée en concurrence sur un marché, une caisse publique ou encore l'Etat. Ils sont néanmoins confrontés à des objectifs et des contraintes communes.

Rémunérations des médecins : quels enjeux pour le financeur ?

Considérons deux cas polaires d'organisation du financement des soins : (i) des assurances en concurrence et (ii) une assurance publique obligatoire ou un système de financement étatique de type beveridgien (voir chapitre ??). Un assureur en situation de concurrence visera naturellement la rentabilité financière de son activité alors que nous ferons l'hypothèse que le financeur public cherchera à produire le service de soins le plus efficace possible compte tenu de son budget. L'assureur privé est en concurrence sur sa clientèle ce qui n'est pas le cas d'un financeur public. Il fait face en outre à un deuxième marché celui des offreurs de soins avec

lesquels il doit contractualiser. Le financeur public est soumis lui aussi, quoique dans une moindre mesure, à cette concurrence sur les offreurs de soins puisque dans de nombreux pays organisé autour d'un financement public des soins l'offre de soins public doit faire face à une concurrence d'une offre de soins privée.

Que la contrainte soit une contrainte principalement concurrentielle (assurances en concurrence) ou principalement budgétaire (financement public) cela conduit les financeurs à partager des objectifs communs. Le premier objectif qui s'impose est celui de **l'efficience** c'est-à-dire celui de la parcimonie : le moindre coût pour un service de soins donné. Sous l'hypothèse d'un marché effectivement concurrentiel de l'assurance, un assureur privé est contraint de proposer une prime d'assurance la plus faible possible pour un panier de services de soins donné afin d'être compétitif sur le marché. Le financeur public est quant à lui tenu par ses contraintes de financement (contraintes de budgets publics financés par impôts ou cotisations sociales) et plus généralement par des objectifs de bonne gestion des fonds publics. Cet objectif d'efficience entre en tension avec les objectifs des professionnels de santé qui sont attentifs à leur niveau de rémunération mais aussi à leur qualité de vie au travail (temps de travail, charges administratives, flexibilité des horaires, qualité de l'interaction avec les patients ...). Les négociations sur les rémunérations entre financeurs et professionnels de santé organisent le partage de la valeur produite par les soins entre le financeur (et ses clients ou ses administrés) et les professionnels de santé. La tension entre les professionnels de santé et le financeur est une réalité de fait. Les professionnels tendent à souhaiter bénéficier de l'ensemble de la valeur qu'ils créent (c'est-à-dire bénéficier de la rémunération la plus élevée possible au regard des dispositions à payer des usagers du système de santé) alors que le financeur souhaite que le partage de cette valeur soit plus avantageux pour les usagers (pour être concurrentiel sur le marché pour l'assureur privé ou pour limiter les dépenses pour le financeur public).

Le second objectif qui s'impose au financeur est celui de l'attractivité des professionnels de santé. Pour un assureur privé c'est un enjeu majeur puisqu'il est en concurrence avec les autres assureurs (même si un offreur de soins donné peut généralement contractualiser avec plusieurs assureurs). Pour s'assurer une taille de marché importante il doit contractualiser avec un grand nombre d'offeurs de soins afin de proposer un volume de services de soins suffisant pour ses clients. Pour un assureur la rémunération qu'ils proposent aux médecins est un instrument déterminant pour attirer et conserver les offreurs de soins dans le panier de services offert à ses assurés. Dans les pays dont le financement des soins s'organise principalement autour d'un marché concurrentiel de l'assurance santé le pouvoir de marché des médecins, fonction de leur nombre et de la façon dont ils s'organisent pour contractualiser avec les assureurs, est un paramètre déterminant de leur niveau des rémunérations.

Pour le financeur public la question de l'**attractivité** des médecins se pose de manière moins aiguë mais elle se pose néanmoins. D'une part, il doit s'assurer que les médecins ne se déplacent pas de manière trop massive vers les secteurs d'activité dans lesquels les prix ne sont pas administrés (le secteur avec dépassement d'honoraires en France ou le secteur privé en Allemagne par exemple) pour garantir l'accès aux soins de l'ensemble de la population et de la population la plus défavorisée en particulier. Le cas français est particulièrement intéressant de ce point de vue. L'ouverture en 1980 d'un secteur à honoraire libre a été rapidement conditionné restreint, dès 1989, aux chefs de cliniques pour éviter que les difficultés financières d'accès aux soins prennent trop d'ampleur (Bras, 2015). Le secteur 2 est maintenant conçu comme un incitatif financier pour que les internes prolongent leur activité à l'hôpital avant d'exercer en ville. Cette opportunité laissée aux jeunes spécialistes posent maintenant des questions d'accessibilité aux soins de spécialistes pour les personnes les plus modestes résidant dans des zones géographiques dans lesquelles une minorité de médecins spécialistes ne pratiquent pas de dépassement d'honoraires. Les inégalités d'accès aux soins de spécialistes sont d'ailleurs particulièrement importantes en France comparativement à nos voisins européens (voir chapitre **???**) même si d'autres obstacles que des obstacles financiers permettent d'expliquer cette réalité. C'est ainsi que l'assurance maladie propose aujourd'hui aux médecins exerçant en secteur 2 des contrats incitatifs visant à réguler le montant global des dépassements d'honoraires (Optam). Le principe du contrat est d'inciter les médecins en secteur 2 à modérer leur dépassement d'honoraires. Cet exemple montre la difficulté pour le financeur public de négocier des modèles de rémunération qui garantissent les différents objectifs d'attractivité qu'il peut légitimement se fixer : ici (i) maintenir dans l'hôpital public les jeunes médecins et (ii) garantir un égal accès aux soins des spécialistes en conservant une part la plus large possible de l'activité des médecins de ville dans le cadre des prix administrés.

D'autre part, plus généralement, le financeur public doit favoriser une attractivité suffisante du métier de médecin pour garantir la formation d'un nombre suffisant de médecins capables d'offrir des services de santé de qualité. L'attractivité des professions dont les rémunérations sont encadrées (comme dans la fonction publique) est une question de première importance pour les systèmes de santé des pays à revenus faibles ou modérés, c'est également un enjeu important de long terme pour la qualité des systèmes de santé dans les pays les plus riches.

S'ajoute aux enjeux d'efficacité et d'attractivité celui de la **qualité des soins** que l'assureur soit privé et en concurrence ou un acteur public : le premier par nécessité concurrentielle, le deuxième pas soucieux de bonne gestion. Il y a cependant deux manières d'apprécier la qualité des soins. Elle peut s'apprécier par la capacité des soins délivrés à satisfaire les patients ou par la capacité des soins à contribuer à la santé des patients. Bien entendu, ces deux façons de

définir la qualité des soins peuvent se rejoindre mais elles ne sont pas toujours parfaitement superposables car les patients n'ont qu'une information imparfaite (voir chapitre ??) sur la capacité des soins délivrés à favoriser leur santé sur le court et le long terme. La concurrence entre assureurs de soins, quand elle ne se traduit pas une concurrence entre leurs offres de services de soins, organisent une concurrence entre offreurs de soins et peut ainsi favoriser la qualité comprise comme la satisfaction des patients qui expriment leurs préférences par le choix de leur assureur. Mais la concurrence entre offreurs de soins peut également s'organiser quand le financement est assuré par un financeur public dès lors que les usagers du système de santé ont la possibilité de choisir leur médecin. C'est bien le cas en France où l'existence d'une assurance maladie obligatoire publique autorise les usagers du système de santé à choisir leur médecin et à ainsi les mettre en concurrence. Si les mécanismes concurrentiels sont des instruments de régulation qui peuvent favoriser la qualité des soins cela n'interdit pas de s'interroger sur les modalités de rémunération qui favorisent la qualité des soins. Cette question est d'autant plus pertinente quand on peut douter de l'efficacité des mécanismes concurrentiels (asymétrie d'information entre professionnels de santé et patients, alternatives d'offres de soins insuffisantes). Nous reviendrons sur ce point dans les sections suivantes.

La qualité des soins doit également être mesurée au regard de l'accessibilité des soins sur l'ensemble du territoire dans des délais d'attente raisonnables ne mettant pas en jeu la santé des patients. Les effets des modalités de rémunération sur le temps de travail des médecins et le choix de leurs lieux d'exercice est un enjeu d'importance pour la régulation du système de santé considérant que le nombre de professionnels dans un système de santé est fixe à court terme et ne peut évoluer que sur le moyen/long terme en raison de la durée de formation d'un médecin.

Après avoir discuté des conditions dans lesquelles les rémunérations des médecins étaient négociées avec les financeurs il est temps de discuter des principales modalités de rémunération que les financeurs peuvent négocier avec les offreurs de soins.

Les différentes modalités de rémunérations des médecins et leurs propriétés

On rencontre dans la plupart des systèmes de soins deux grands types de mode de rémunération des médecins et professionnels de santé exerçant en ville.

Le paiement à l'acte

Le mode de rémunération qui est très largement dominant en France (mais aussi dans de nombreux autres pays, en particulier dans les pays dont le système de santé est d'inspiration bismarckienne) est celui du paiement à l'acte. Le professionnel tarifie un prix pour chaque acte, consultation comprise. Il relève d'un principe marchand usuel, le client (patient) paye pour chaque service offert (en anglais on parle ainsi de « fee for service »). Ce mode de rémunération est profondément ancré dans l'histoire du système de santé français, et ce, bien avant la création de l'assurance maladie en 1945 : les médecins exercent une activité libérale, ils tarifient leur activité (en fonctions des services offerts) comme toute profession libérale. L'assurance maladie est une institution qui vient couvrir les patients contre le risque financier lié aux soins mais cette couverture assurantielle ne doit pas venir contrarier le lien entre le médecin et son patient qui relève d'un échange marchand. C'est une des raisons pour lesquelles certains syndicats de médecins se sont élevés contre le tiers payant généralisé (qui devait se traduire par l'absence d'avance de frais pour les patients) qui rompait ce lien marchand et risquait d'assujettir directement et entièrement la rémunération des médecins aux paiements par l'assurance maladie des services de soins. Cette vision syndicale historique de l'activité libérale est toujours très ancrée chez les médecins libéraux même si elle aujourd'hui contestée par une partie des professionnels.

Au-delà des considérations politiques attachées à la rémunération à l'acte, qui constituent néanmoins une contrainte politique forte pour le régulateur, ce mode de rémunération à des caractéristiques spécifiques en termes d'incitatifs économiques :

1. C'est un mode de rémunération qui est incitatif à l'activité. Le médecin est rémunéré pour chaque acte réalisé. Son revenu est ainsi directement fonction de son volume d'activité (Hickson et al, 1987). C'est un atout pour le régulateur quand la densité de médecins est faible et que la question première pour le régulateur est de garantir l'accès aux soins à tous.
2. Il a en revanche le désavantage de favoriser la demande induite (voir chapitre ???). Le risque de surconsommation est d'autant plus important que la densité de médecins est importante c'est-à-dire que la concurrence est forte. Comme nous l'avons discuté dans le chapitre ??? la demande induite peut néanmoins jouer sur le type d'acte réalisé et non pas sur le volume et générer une perte d'efficacité même dans un contexte où la densité de médecins est faible.
3. C'est un mode de rémunération qui s'inscrit dans un cadre d'activité libérale et donc dans un contexte de concurrence sur la patientèle (tout au moins quand l'offre de soins est suffisante pour rendre effective la concurrence). Il favorise donc la qualité des soins ou plus précisément la qualité des soins perçue par les patients.
4. C'est un mode de rémunération qui ne favorise pas les consultations longues la rémunération d'une consultation n'étant pas attachée à sa durée, en tous les cas en France. C'est un mode de rémunération qui est ainsi considéré comme peu favorable aux

consultations de prévention, nécessitant des temps longs d'échanges avec le patient. Certains pays (Canada, Belgique, Suisse, Pays-Bas, Allemagne ...) ont instauré des tarifs de consultations variables en fonction du type de consultation (motif et type de patient) et de sa durée afin de limiter l'effet délétère du paiement à l'acte sur la durée des consultations. La contrepartie de ces ajustements est la gestion d'un système d'information permettant au financeur d'identifier chaque type de consultation ce qui peut générer une charge administrative supplémentaire pour les médecins devant rendre compte de leur activité.

On oppose à la rémunération à l'acte la rémunération à la capitation.

La capitation (paiement par tête)

La capitation consiste à allouer au médecin une somme d'argent (généralement annuel) pour chaque patient de sa patientèle. Cela suppose donc que chaque usager du système de soins soit rattaché à un médecin (le médecin traitant en France). Ce mode de rémunération concerne généralement les médecins généralistes. Le montant de la capitation peut varier en fonction des caractéristiques des patients (en fonction des besoins de soins attendus du patient).

L'exemple emblématique est celui de la Grande-Bretagne dont le système de soins repose sur des médecins généralistes qui se présentent comme la porte d'entrée dans le système de soins (« Gatekeepers ») et qui sont essentiellement rémunérés à la capitation.

Ce système de rémunération à des propriétés opposées de celles du paiement à l'acte.

1. C'est un système qui n'est pas incitatif à l'activité puisque la rémunération des médecins est (partiellement au moins) déconnectée de l'activité. C'est un avantage dans le sens où cela permet d'éviter les phénomènes de demande induite étudiés précédemment. Néanmoins cela peut conduire les médecins à adresser plus souvent les patients vers des spécialistes (cela réduit leur activité sans réduire leur revenu) ce qui est une autre forme de demande induite (même si la demande induite passe par l'adressage à un autre médecin).¹ Comme nous l'avons discuté est un désavantage quand la densité des médecins est faible au regard des besoins de la population. Dans cette configuration le régulateur souhaite que les médecins aient un volume d'activité important pour compenser le manque de médecins.
2. C'est un mode de rémunération qui favorise également la concurrence sur la patientèle (dès lors que la densité de médecins est suffisante) puisque les usagers peuvent changer tous les ans de médecins traitants (même si la concurrence est certainement moins vive que dans un système de paiement à l'acte). C'est une des différences importantes avec le salariat qui constitue également une modalité courante (mais très minoritaire en France) de rémunération des médecins de premier recours et dont les propriétés se rapprochent de celle du paiement à la capitation.

¹ Voir Iversen et al, 2000..

3. C'est un mode de rémunération qui ne défavorise pas les consultations longues (de prévention par exemple) puisque la durée des consultations n'a pas d'effet sur la rémunération des médecins.
4. C'est un mode de rémunération qui permet au financeur de contrôler les dépenses de soins puisque les dépenses ne dépendent pas des actes mais seulement du nombre de patients. Le risque d'une demande plus importante est portée par les médecins. Bien sûr on peut imaginer des mécanismes de compensation ex-post en cas de choc sur les besoins de soins (une épidémie par exemple).

Dans la plupart des systèmes de soins ces deux modes de rémunération, capitation et paiement à l'acte, coexistent. En Grande Bretagne, par exemple, certains actes sont tarifés en sus de la capitation pour inciter les médecins à prodiguer plus fréquemment ces actes. En France, pour prendre un exemple de pays dans lequel la rémunération à l'acte est prédominante, des forfaits annuels ont été introduits pour les médecins généralistes avec des majoration pour les patients en Affection de Longue Durée (ALD) dans le but de rendre compte de la surcharge de travail liée au suivi de ces patients. La rémunération des généralistes demeure néanmoins toujours dépendre à plus de 80% du paiement à l'acte.

Nous proposons dans la section suivante une discussion analytique sur la combinaison optimale de ces deux modes de rémunération.

En complément de la capitation et du paiement à l'acte, les financeurs utilisent fréquemment des systèmes de paiement à la performance. On peut en définir deux sortes en fonction du type de performance recherchée

1. Les paiements à la performance visant des objectifs de santé publique et, pour reprendre la discussion précédente, visant la qualité des soins appréciée par l'effet sur la santé des patients des services de soins. Ce type de paiement fondé sur des indicateurs de performance existe aussi bien en Grande Bretagne, dans un système à capitation, qu'en France, dans un système avec paiement à l'acte. Ces indicateurs peuvent porter sur des objectifs en termes de taux de vaccination, de dépistage, suivi des patients souffrants d'une affection chronique Le médecin reçoit en fin d'année un paiement fonction du degré de réalisation des objectifs fixés par le financeur.²
2. Les paiements à la performance peuvent aussi se fixer comme objectif une utilisation parcimonieuse des ressources. Les médecins sont alors intéressés à des objectifs en termes de dépenses. Si les dépenses de soins pour une patientèle donnée sont inférieures à un objectif fixé à l'avance le médecin est intéressé à la différence entre l'objectif de dépenses et les dépenses effectives. C'est un moyen pour le financeur d'inciter financièrement les médecins à l'usage parcimonieux des ressources. Le plus souvent ce type de paiement n'est pas défini pour un médecin isolé mais pour une équipe de soins intervenant auprès d'une catégorie donnée de patient.

² Voir Hurley et al, 2014 pour les résultats contrastés de l'évaluation de l'impact d'un système de paiement à la performance au Canada.

3. La performance, tant du point de vue de la qualité des soins que de l'usage parcimonieux des ressources, est souvent dépendant d'une prise en charge coordonnée des patients par l'ensemble professionnels de santé concernés par la prise en charge (infirmiers, médecins généralistes, kinésithérapeute, spécialistes, services hospitaliers ...). La coordination des professionnels de santé est particulièrement importante pour les patients souffrant d'une maladie chronique (et cumulant parfois plusieurs pathologies au long cours) nécessitant l'intervention de plusieurs professionnels de santé (référence). Cette question de la coordination des soins est en enjeu important des systèmes de santé des pays vieillissants. Les réformes visant à favoriser la coordination entre les différents professionnels de soins sont généralement dénommées politiques **d'intégration des soins**. On parle en économie industrielle de logique d'intégration de filière quand les coûts de coordination (ou d'absence de coordination) d'entités de productions autonomes sont trop importants.

Un des instruments pour favoriser cette intégration est de proposer des paiements forfaitaires non pas professionnels par professionnels mais à l'équipe de soins prenant en charge le patient. Ceci peut s'apparenter à une capitation collective, une capitation à l'équipe de soins. On parle de paiement groupé (bundle payment) ou paiement à l'épisode quand il s'agit de soins qui font suite à un événement aigu (une prothèse de hanche par exemple). Ce mode de paiement demande une forte intégration des soins puisque cela suppose d'organiser la répartition du montant de la capitation entre les différents membres de l'équipe de soins (voir Yordanov et al, 2024). Le plus souvent la mission est confiée à un des professionnels (ou collectif de professionnels comme un service hospitalier pour prendre un exemple fréquent) de l'équipe de soins. Ce type de paiement peut permettre d'intéresser financièrement l'équipe de soins à la gestion parcimonieuse des ressources grâce à des bonus dépendant de la réalisation d'objectifs en termes de dépenses comme nous l'avons mentionné plus haut. Cela peut aussi permettre de fixer des bonus financiers en fonction de la qualité des soins en définissant des objectifs en termes d'indicateurs de santé pour la patientèle prise en charge par l'équipe de soins.

Comme nous venons de le discuter les objectifs d'efficience, de qualité poursuivis par les financeurs peuvent se traduire par des modalités de rémunération très variées en fonction des systèmes de santé et d'organisation du financement des soins. On perçoit la difficulté pour le financeur de définir des systèmes de paiement optimaux, que le contexte de financement des soins soit concurrentiel ou non, compte tenu de la complexité des systèmes de santé.

Les travaux empiriques sur les effets de ces différents modes de rémunération sur l'activité des médecins sont nombreuses mais les résultats sont contrastés et d'un niveau de preuve souvent assez faible au regard des standards de « l'evidence based medicine » (Jia et al, 2021). La nature contrastée des résultats tient au fait de l'importance du fonctionnement global du système de santé sur l'effet de la modification des modes de rémunération des médecins sur leur activité. Les facteurs influençant ces résultats sont très nombreux et à chaque fois différents en fonction des contextes nationaux ou régionaux. Les niveaux de preuve sont la plupart du temps

modestes en raison de la difficulté méthodologique de construire des designs d'études reposant sur une randomisation suffisante. Il s'agit le plus souvent d'évaluer ex-post des réformes et non de construire *a priori* des expérimentations randomisées. Ceci ne veut pas dire pour autant que les modes de rémunérations n'ont pas d'effet sur l'activité des médecins mais seulement que ces effets doivent être contextualisés et évalués avec la prudence requise au regard des difficultés d'identification empirique de leurs amplitudes voir de leurs sens

Nous proposons dans la section suivante de simplifier le cadre d'étude pour conduire une discussion analytique plus approfondie permettant de mieux comprendre les complémentarités offertes en théorie par les paiements à l'acte et à la capitation.

Discussion sur la tarification optimale

Nous avons précédemment discuté des objectifs et des contraintes qui s'imposaient aux financeurs dans leurs négociations avec les professionnels de santé sur le niveau et les modalités de leurs rémunérations. Nous avons également décrit les deux principaux types de modalités de rémunération : capitation et paiement à l'acte. L'objet de cette dernière section est de discuter du mode de tarification optimale combinant ces deux modes de rémunération dans un cadre simplifié avec un seul assureur couvrant entièrement les dépenses de soins des patients.

Nous rappelons qu'une tarification optimale (au sens de Pareto) pour un économiste signifie qu'il n'existe pas une autre tarification qui permettrait à tous, ici médecins et patients, d'y gagner. C'est une forme d'optimalité sociale.

Pour conduire cette discussion nous nous appuyons sur le cadre formel de McGuire (2011). On considère que la quantité de soins, x , pour un patient est fixée par le médecin. Nous faisons ainsi l'hypothèse que la demande est entièrement induite par le médecin. Pour ce patient le médecin reçoit du financeur une capitation, F , c'est-à-dire une somme d'argent indépendante de la quantité de soins délivrée, et un montant fx proportionnel à la quantité de soins où f désigne le prix unitaire (le prix de l'acte, la consultation par exemple). Le bénéfice pour le médecin s'écrit :

$$\pi(x) = F + fx - cx$$

où c désigne le coût marginal du service de soins pour le médecin. Ce coût recouvre les coûts liés à l'exercice médical lui-même mais il faut également y ajouter le coût d'opportunité c'est-à-dire le coût subjectif occasionné par le temps de travail nécessaire à la réalisation des soins c'est-à-dire le coût subjectif lié à la diminution du temps de loisir correspondant. On fera ici l'hypothèse que le coût marginal est constant à l'instar de McGuire (2011). C'est une

simplification très forte parce que l'on peut légitimement présumer que le coût d'opportunité augmente avec x c'est-à-dire avec le temps de travail. Elle se justifie à l'échelle d'un patient car le temps de travail reste nécessairement réduit à cette échelle, elle se justifie plus difficilement à l'échelle de la patientèle du médecin. Nous reviendrons plus loin sur les enjeux de cette simplification. L'ajout d'un coût fixe (indépendant de x) serait légitime pour rendre compte des coûts fixes liés à l'exploitation d'un cabinet médical mais cela ne changerait pas les conclusions de l'analyse que nous allons conduire.

Comme dans le chapitre ??? nous faisons l'hypothèse que le comportement du médecin peut se comprendre par la maximisation d'une fonction d'utilité rendant compte des motivations intrinsèques et extrinsèques du médecin :

$$V(x) = V[B(x), \pi(x)]$$

$B(x)$ désigne le bénéfice pour le patient tiré de la quantité x de soins que l'on suppose croissant avec x ($B'(x) \geq 0$) mais avec des rendements décroissants ($B''(x) < 0$: le gain marginal d'une augmentation de x décroît avec x). On rappelle que le patient est supposé couvert à 100% par son assurance et que le bénéfice est fonction des effets (perçus) des soins sur sa santé. On suppose naturellement que l'utilité du médecin est fonction croissante du bénéfice pour le patient (motivations intrinsèques, $V'_B > 0$) et fonction croissante de son propre bénéfice (motivations extrinsèques, $V'_\pi \geq 0$)

On supposera que ce bénéfice pour la santé s'exprime en disposition à payer c'est-à-dire en argent. La quantité optimale de soins x^* pour un patient donné est caractérisée par la condition d'optimalité : $B'(x) = c$. Il est socialement optimal de continuer d'augmenter la quantité de soins délivrée tant que le bénéfice marginal du soin est supérieur au coût. Il est par exemple socialement optimal de permettre une consultation supplémentaire tant que la disposition à payer du patient est supérieur au prix de la consultation (en assimilant ici c au prix de la consultation). C'est le raisonnement d'équilibre sur un marché que nous avons déjà discuté dans le chapitre ???.

Plaçons-nous maintenant du point de vue du médecin qui est supposé décider de la quantité x de soins qu'il va délivrer au patient considéré. Commençons par supposer que le médecin est parfaitement bienveillant et que sa fonction d'utilité ne soit fonction que du bénéfice des soins pour le patient ($V'_\pi = 0$) :

$$V(x) = V[B(x)]$$

Dans ce cas le médecin va augmenter la quantité de soins tant que le patient en tire un bénéfice, c'est-à-dire jusqu'à la quantité de soins x^b telle que : $B'(x) = 0$.

Comme c'est illustré sur le graphique 1 la quantité de soins va être à un niveau supérieur au niveau socialement optimal. Ceci vient du fait que nous supposons ici que le coût des soins n'est pas prise en compte dans la décision du médecin. C'est le mécanisme de l'aléa moral discuté dans le chapitre ???.

Faisons maintenant l'hypothèse que le médecin est imparfaitement bienveillant et considère également son bénéfice dans sa prise de décision ($V'_\pi > 0$). Dans ce cas la condition d'optimalité est plus compliquée puisque le médecin maximise la fonction d'utilité définie par l'identité (1) ce qui conduit à considérer les effets de la quantité de soins sur le bénéfice du patient mais aussi sur le bénéfice du médecin. La condition d'optimalité s'écrit ainsi : $B'(x)V'_B(x) + \pi'(x)V'_\pi(x) = 0$. On peut réécrire cette identité ainsi :

$$B'(x) = \frac{\pi'(x)V'_\pi(x)}{V'_B(x)}$$

Sachant que $\pi'(x) = f - c$, on peut considérer deux cas (notons que le cas $f - c = 0$ est équivalent au cas d'un médecin parfaitement bienveillant ($V'_\pi = 0$)) :

1. $f - c > 0$; dans ce cas $\frac{\pi'(x)V'_\pi(x)}{V'_B(x)} < 0$ et la quantité produite x^+ sera supérieure à la quantité produite par un médecin parfaitement bienveillant ($V'_\pi = 0$). Se rajoute à l'aléa moral généré par la demande du patient un aléa moral provoqué par un effet de demande induite positif (voir graphique 1 sous l'hypothèse $\frac{V(x)}{V'_B(x)}$ constant).
2. $f - c < 0$, dans ce cas $\frac{\pi'(x)V'_\pi(x)}{V'_B(x)} > 0$ et la quantité produite x^- sera inférieure à la quantité produite par un médecin parfaitement bienveillant ($V'_\pi = 0$). L'effet de demande induite (négatif puisque le prix net des soins est négatif) vient tempérer l'aléa moral général par la demande du patient (voir graphique 1 sous l'hypothèse $\frac{V(x)}{V'_B(x)}$ constant).

Dans ce dernier cas, on constate qu'il est théoriquement possible pour le financeur de fixer f suffisamment bas de telle sorte que la quantité produite soit socialement optimale (soit f de telle sorte que $\frac{\pi'(x)V'_\pi(x)}{V'_B(x)} = c$ sous l'hypothèse $\frac{V(x)}{V'_B(x)}$ constant). Bien sûr dans ce cas le médecin peut être confronté à un bénéfice d'activité négatif ce qui peut remettre en question son activité de soins, sa participation même au système de santé. Le financeur peut néanmoins utiliser le deuxième instrument de la tarification, la capitation F , pour assurer un revenu suffisant au médecin. Pour les économistes le paiement à l'acte f doit être mobilisé pour conduire le médecin à fixer le niveau optimal de soins x alors que la capitation F doit être mobilisée pour fixer le bon niveau de rémunération des médecins (celui qui garantit une attractivité suffisante du métier). Les deux instruments devraient être utilisés ensemble.

Il est important de noter qu'il est attendu des médecins, dans cette logique de paiement optimal, qu'ils fournissent des soins qui ne sont pas rentables pour eux (ils y perdent, $f - c < 0$) en jouant sur leurs motivations intrinsèques. C'est une situation qui entre en tension avec la logique de

l'activité libérale de la médecine fondée sur un échange marchand mutuellement avantageux avec le patient. C'est une des raisons qui expliquent la difficulté d'introduire de manière significative ce type de paiement mixte en France alors que l'exercice de la médecine de ville est historiquement ancrée dans un mode d'activité libérale.

Il faut pour finir garder à l'esprit que dans un contexte dans lequel la densité de médecin est faible au regard des besoins de la population, le coût marginal de production du service de soins peut être fortement croissant avec la quantité de soins produite pour un patient donné et la quantité optimale définie par l'identité (??) être considérée comme trop faible. En quelque sorte, le coût marginal de production peut être considéré comme le reflet d'un niveau d'offre de soins sous-optimal. Le coût marginal ne reflète plus alors le coût marginal de production socialement optimal (c'est-à-dire pour une offre de soins optimale). Dans ce cas le niveau de production défini par l'identité entre coût marginal de production et bénéfice marginal pour le patient ne peut plus être considéré comme optimal. Il est alors possible de justifier un paiement à l'acte avec un prix f élevé, plus élevé que le coût marginal, pour encourager l'activité des médecins.

La définition d'une mode de paiement optimal ne peut être caractérisé sans prendre en considération le niveau de l'offre de soins. Dans le cadre formel de cette section cela revient à s'interroger sur le fait de savoir si le coût marginal reflète le niveau optimal de l'offre de soins.

Enfin, nous avons ici raisonné en considérant que le bénéfice exprimé par le patient, et pris en compte par le médecin, reflétait des préférences éclairées (connaissances éclairées de l'effet des soins prodigués sur sa santé) ce qui peut être contesté en particulier dans un contexte de concurrence sur la patientèle qui peut conduire les médecins à satisfaire des préférences qui ne sont pas nécessairement uniquement en lien avec l'effet des soins sur la santé.

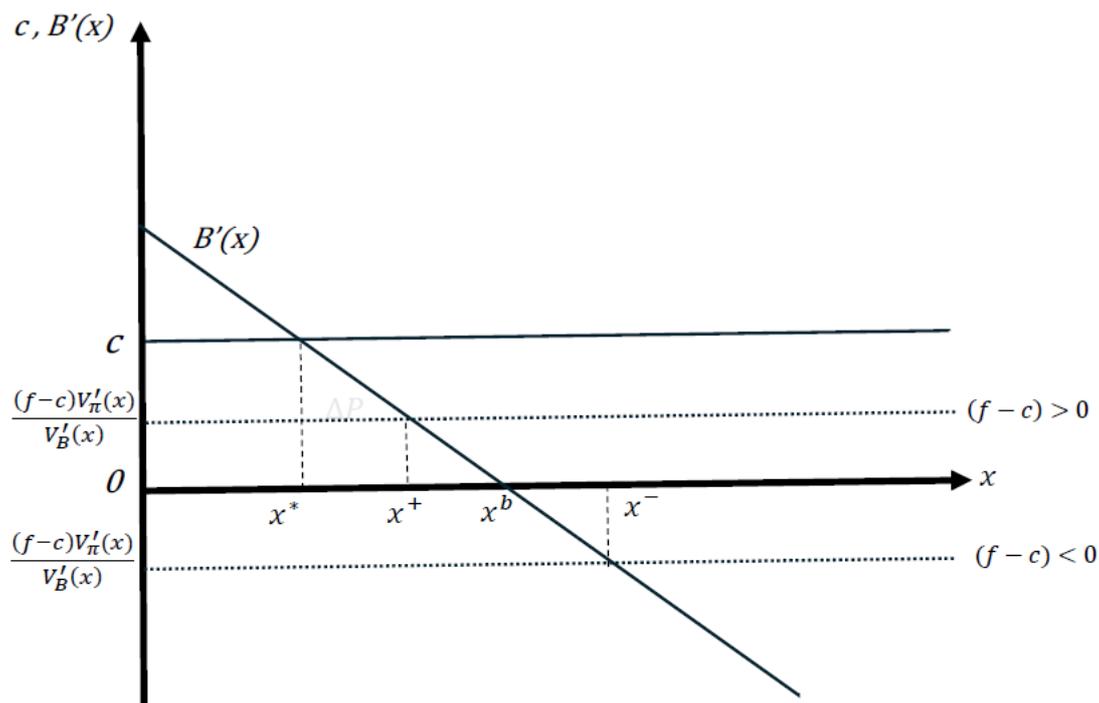


Figure 1 : tarification optimale

Références

Bras P-L, 2015, La liberté des tarifs médicaux : la victoire de médecins spécialistes, Les Tribunes de la Santé, n°48.

Hickson_GB, Altemeier_WA, Perrin_JM. Physician reimbursement by salary or fee-for-service: effect on physician practice behavior in a randomized prospective study. *Pediatrics* 1987;**80**(3):344-50.

Iversen T, Lurås H. The effect of capitation on GPs' referral decisions. *Health Econ.* 2000 Apr;**9**(3):199-210.

Jia L, Meng Q, Scott A, Yuan B, Zhang L. Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database of Systematic Reviews* 2021, Issue 1. Art. No.: CD011865.

Li_J, Hurley_J, DeCicca_P, Buckley_G. Physician response to pay-for-performance: evidence from a natural experiment. *Health Economics* 2014;**23**(8):962-78.

McGuire TG. Physician agency and payment for primary medical care. In The Oxford Handbook of Health Economics, S. Glied, P.C. Smith (eds). Oxford University Press, 2011; 602-623.

Dimitar Yordanov, Anne Sophie Oxholm, Thim Prætorius, Søren Rud Kristensen 2024. Financial incentives for integrated care: A scoping review and lessons for evidence-based design, Health Policy, Volume 141.

Comment citer ce cours ?

Economie de la santé, Jérôme Wittwer, AUNEGe (<http://aunege.fr>), CC – BY NC ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Cette œuvre est mise à disposition dans le respect de la législation française protégeant le droit d'auteur, selon les termes du contrat de licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). En cas de conflit entre la législation française et les termes de ce contrat de licence, la clause non conforme à la législation française est réputée non écrite. Si la clause constitue un élément déterminant de l'engagement des parties ou de l'une d'elles, sa nullité emporte celle du contrat de licence tout entier.

Figures

Figure 1 : tarification optimale 15