

**Principes et enjeux de la tarification
à l'activité à l'hôpital (T2A)
Enseignements de la théorie économique
et des expériences étrangères**

Zeynep Or (IRDES)
Thomas Renaud (IRDES)

DT n° 23

Mars 2009

Principes et enjeux de la tarification à l'activité à l'hôpital (T2A)

Enseignements de la théorie économique et des expériences étrangères

Zeynep Or, Thomas Renaud

Résumé

La France a introduit en 2005 un système de tarification à l'activité (T2A) pour financer les établissements de santé, comme une vingtaine de pays avant elle. Malgré des principes de bases communs, les systèmes de T2A varient néanmoins dans leur architecture. Afin de mieux comprendre la problématique française de la T2A, ce document de travail examine les justifications économiques d'un tel système, les dispositifs-clés dans sa mise en œuvre et les enjeux qui en découlent.

Le principe de base de la T2A, qui consiste à payer les établissements en fonction de leur activité mesurée par groupe homogène de malades, a des avantages pour améliorer l'efficacité et la transparence dans le financement des soins. Toutefois, la littérature et les expériences des autres pays présentés dans ce papier montrent que ce mécanisme de paiement présente un certain nombre de risques et nécessite des ajustements réguliers et soigneux pour obtenir les bénéfices attendus d'un tel système. Assurer la cohérence à la fois clinique et économique du classement de l'activité hospitalière, et établir le niveau des tarifs correspondant constituent deux véritables défis. Le principe de payer un prix fixe qui soit directement indexé sur les coûts moyens observés et qui reste commun à tous les types d'établissements est de plus en plus contesté.

Par ailleurs, la T2A, par sa nature, peut induire des effets pervers : il est donc nécessaire de prévoir des mécanismes régulateurs complémentaires pour garantir la qualité des soins et un accès équitable. Dans une perspective de maîtrise des dépenses de santé, il est également important de suivre de près l'évolution de l'activité dans les différents modes de prise en charge hospitalière, ainsi qu'en médecine de ville, puisque la T2A incite les établissements à augmenter leur activité en induisant la demande de soins et à transférer une partie de leurs coûts vers les soins de suite ou de domicile.

Mots clés : tarification à l'activité, hôpital, régulation, comparaison internationale

Code JEL : H3, H4.

1. Introduction

Depuis 2005¹, l'activité de court séjour des hôpitaux publics et privés est financée par la tarification à l'activité (T2A). Cette nouvelle tarification a pour but de fonder le financement des établissements en fonction de leur activité, celle-ci étant décrite par groupe homogène de séjours (GHS). Elle couvre 100 % de l'activité des établissements privés depuis 2005, tandis qu'elle a été appliquée de manière progressive dans le secteur public pour couvrir 100 % de l'activité de court séjour (Médecine, Chirurgie, Obstétrique) des établissements publics en 2008.

La France n'est pas le premier pays à introduire une tarification à l'activité des soins hospitaliers. En adoptant ce mode de paiement elle rejoint une vingtaine de pays qui ont déjà mis en place une tarification de ce type. Malgré des principes de base communs, ces systèmes de tarification à l'activité varient néanmoins dans leur architecture globale et chaque pays adopte les paramètres de fixation des tarifs en fonction de ses besoins et contraintes propres. Il n'existe donc pas un « design » unique de T2A.

Cependant la structure du système de tarification, la précision et l'homogénéité des classifications d'activité utilisées et les modalités techniques de fixation des tarifs déterminent directement les incitations pour les différents acteurs et la réussite des objectifs politiques.

Afin de mieux comprendre la problématique française de la T2A, ce papier examine les justifications économiques d'un tel système, les mécanismes/dispositifs clés dans sa mise en œuvre et les enjeux qui en découlent.

Pour cela nous introduisons tout d'abord les principes économiques constitutifs d'une tarification à l'activité et identifions les variables structurelles qui modulent directement son impact (sections 2.1 et 2.2). Nous évoquons ensuite ses avantages (section 2.3) mais aussi ses effets pervers potentiels sur la qualité des soins (section 3.1) et la maîtrise des dépenses (section 3.2). Les principaux défis communs à tous les systèmes de tarification à l'activité sont développés dans la section 4, avec un zoom particulier sur la problématique de fixation des tarifs. Enfin, les enseignements des expériences à l'étranger sont exposés dans la section suivante (section 5) pour enrichir notre réflexion, expliquer comment la mise en œuvre peut varier d'un pays à l'autre et rappeler les enjeux qui en découlent pour les décideurs publics.

Cette analyse est complétée par ailleurs, par une réflexion plus ciblée sur la convergence des tarifs hospitaliers publics/privés et sur les sources de variation des coûts de prise en charge hospitalière, étayée par une analyse empirique de l'activité hospitalière française (Or *et al.* 2009).

¹ La loi de financement de la sécurité sociale pour 2004.

2. Les bases économiques d'une tarification à l'activité

Après avoir défini ce que l'on entend par Tarification à l'activité (T2A) et expliqué le fonctionnement général d'un tel système et les conditions de sa mise en place, nous clarifions les effets possibles de son application. Nous nous aidons d'exemples étrangers pour illustrer les différents modèles de T2A existants et leurs effets, afin de mieux comprendre les répercussions possibles du modèle employé en France.

2.1. Qu'est qu'une T2A ?

La T2A est un système de financement des établissements de santé qui associe le paiement à l'activité réalisée, celle-ci étant définie par des épisodes de soins.

Deux éléments fondamentaux président à la T2A.

Premièrement, l'activité hospitalière est définie et décrite à travers des groupes homogènes de malades (GHM) plutôt que selon les disciplines de services hospitaliers (ou spécialités). Par exemple, l'établissement reçoit un paiement pour un patient à qui l'on a posé une prothèse de la hanche plutôt que pour un patient soigné dans le service d'orthopédie. Les GHM, comme leur nom l'indique, identifient les différentes prestations de soins offertes à un même profil de patients. Le principe de base est de payer le même prix pour les mêmes prestations, à condition que l'on puisse fournir une description clinique correcte des patients pris en charge et des différentes prestations délivrées par les établissements de santé.

Deuxièmement, les prix des GHM² sont définis à l'avance (paiements prospectifs). Ces prix peuvent être fixés au niveau national, comme c'est le cas en France, ou au niveau local. Il existe différents principes et mécanismes pour fixer les prix (voir ci-dessous).

Malgré ces éléments communs à tout système de T2A, la mise en application de ces deux principes de base (définition des GHM et fixation des prix) varie largement d'un système à l'autre, ce qui a des répercussions importantes sur l'efficacité individuelle des établissements ainsi que sur le fonctionnement du marché hospitalier dans sa globalité.

2.2. Comment fonctionne un système de T2A ?

Selon les objectifs visés, le fonctionnement d'une T2A intègre des paramètres plus ou moins nombreux et imbriqués.

² En France, le terme GHS (Groupes Homogènes de Séjours) est employé pour désigner les prix des GHM.

Les revenus des établissements de santé (R_E) dépendent donc à la base directement de leur activité. Dans sa forme la plus simple, le revenu de l'hôpital sous T2A augmente de manière linéaire suivant le nombre de patients traités dans chaque GHM (Q) multiplié par le prix (P) de chaque GHM (i).

$$R_E = \sum_i (Q_i \times P_i)$$

Mais en réalité, le revenu hospitalier n'est jamais déterminé uniquement par le nombre de patients soignés. Dans tous les pays où la T2A est introduite, y compris la France, les hôpitaux perçoivent également des recettes (revenus) sous d'autres formes : par exemple pour leurs activités de recherche et d'enseignement, ainsi que pour couvrir les coûts fixes liés à la fourniture de certains services de soins (les services d'urgences, par exemple) ou les coûts liés à des contraintes et particularités locales (aspects géographiques et sociodémographiques de la zone, difficultés d'accès aux soins...). La composition de ces « autres » revenus varie d'un système à l'autre et parfois dans le temps. Si l'on nomme cette source de revenu complémentaire Z , la fonction de revenu d'un établissement devient :

$$R_E = \sum_i (Q_i \times P_i) + Z_E$$

Une tarification à l'activité peut également chercher à modifier les incitatifs financiers (*i.e.* les prix) en fonction de l'activité de production. Dans la plupart des systèmes Européens, la T2A est introduite pour stimuler/augmenter l'activité hospitalière afin de réduire les délais d'attente, notamment dans le cas de la chirurgie programmée. Toutefois, on peut effectivement vouloir augmenter l'activité de manière « contrôlée » du fait de contraintes budgétaires au niveau macro-économique. Il faut également pouvoir éviter que les hôpitaux induisent eux-mêmes la demande pour certains types de soins « profitables » et produisent *in fine* une part de soins « inutiles ».

Ceci est la base des contrats de type volume-prix spécifié par GHM (Street *et al.* 2007). Il y a deux paramètres clés pour ce type d'accord : d'abord, il faut définir un niveau d'activité 'cible' (T_i) pour chaque établissement. En général, cette cible correspond à l'activité historique, mais ceci peut varier en fonction des 'besoins'. Ensuite, il faut décider quel serait le niveau de paiement au-delà de l'activité cible : il correspond en général à une proportion (α) du prix établi pour l'activité cible. Formellement, on peut alors décrire la fonction de revenu comme suit :

$$R_E = \sum_i [(T_i \times P_i) + (Q_i - T_i) \times \alpha \cdot P_i] + Z_E$$

Cette méthode est simple à mettre en pratique dans les systèmes où il y a un seul payeur mais plus difficile quand il existe de nombreux financeurs pour un même hôpital, comme c'est le cas aux États-Unis.

Notons que la philosophie de la régulation prix-volume choisie en France ne correspond pas du tout à la logique économique des contrats volume-prix présentés ici. En France, la maîtrise globale des dépenses hospitalières est assurée, dans le schéma actuel, par les objectifs de dépenses pour les hôpitaux de court séjour (public et privé séparément) qui sont définis à partir de l'ONDAM (Objectif National de Dépenses d'Assurance maladie). Le mécanisme choisi, qui est propre à la France, prévoit une baisse des tarifs en cas d'augmentation de l'activité hospitalière globale et non en fonction des évolutions d'activité de chaque établissement. Ce dispositif, qui ne fait pas de distinction entre les différentes activités produites et qui ne prend pas en compte l'effort individuel des établissements, est problématique et peut engendrer des effets pervers. Ceci génère un système extrêmement opaque pour les établissements avec des évolutions peu prévisibles du marché. De plus, à niveau et gamme d'activité équivalente, un établissement peut se voir « sanctionné » dans son financement, à cause des décisions stratégiques de production des autres établissements.

2.3. Les justificatifs et effets attendus d'une T2A

On peut citer trois avantages principaux de la T2A :

- La T2A vise à améliorer la transparence : elle assure en effet une plus grande transparence dans le financement des soins hospitaliers en liant le financement à la production des soins.
- Elle est également perçue comme un mécanisme « équitable » dans la mesure où on paie le même prix pour un même service pour tous les fournisseurs de soins. Cette équité dépend toutefois de la fiabilité de la classification de l'activité en groupes tarifaires : il est impératif que cette classification soit suffisamment fine, et les groupes suffisamment homogènes, pour que les établissements qui attirent systématiquement les patients les plus lourds ne soient pas pénalisés. Il faut également bien prendre en compte les facteurs exogènes liés au contexte local et que les établissements ne contrôlent pas, car ils peuvent influencer fortement les coûts.
- La T2A vise également à améliorer l'efficience, à la fois de chaque établissement individuellement et de l'ensemble du marché : elle introduit en effet une forme de compétition stimulant l'efficience dans un contexte où ces pressions compétitives étaient inexistantes jusqu'alors. Ceci suppose toutefois que les prix reflètent correctement les coûts des producteurs les plus efficaces.

À ce stade, il faut bien distinguer les objectifs politiques poursuivis à travers une T2A de ceux qu'elle peut réellement atteindre, étant donné les caractéristiques inhérentes à ce mécanisme de financement. Souvent, dans les discours politiques en France comme dans les autres pays, on attribue à la T2A des missions qui sont loin de sa vocation, telle que celle d'améliorer la qualité, l'équité d'accès et la couverture optimale des soins.

En fait, la T2A comme mode de paiement n'a aucune vocation à assurer une couverture optimale des besoins ni à améliorer la qualité des soins. Au contraire, il est nécessaire de prévoir des mécanismes régulateurs additionnels pour garantir la qualité des soins et un accès territorial équitable, comme cela a été fait aux États-Unis (voir ci-dessous, section 5.1). D'autre part, les incitatifs fournis par une tarification basée sur les coûts moyens sont théoriquement neutres sur le niveau d'optimalité de la couverture de soins (à condition qu'ils reflètent bien les coûts réels des établissements). La capacité de canaliser les ressources là où il y en a besoin dépend davantage du pouvoir de négociation des acheteurs/financeurs.

3. Les effets pervers possibles de la T2A

L'application d'une T2A peut induire des effets non souhaités au départ. À cet égard, il est particulièrement important de prêter attention aux conséquences de la T2A sur la qualité des soins et sur la maîtrise des dépenses de santé.

3.1. Qualité des soins

Rien dans un système de T2A ne permet de présager de l'amélioration de la qualité des soins. Ceci est d'ailleurs aussi valable dans un système de financement par budget global. Toutefois, la T2A est souvent critiquée puisqu'elle fournit, de fait, des incitations directes à réduire le coût des séjours. Or, toute la question réside dans la légitimité de ces réductions de coût. S'il est possible, et souhaitable, d'y parvenir par le biais d'un fonctionnement plus efficient, une autre stratégie possible consiste à réduire les soins fournis aux patients, ce qui peut représenter une manière « d'économiser » aux dépens de la qualité.

On peut compromettre la qualité de soins de différentes manières mais, trois stratégies ont été bien débattues dans la littérature.

Premièrement, les établissements ayant intérêt à raccourcir les séjours, ils peuvent renvoyer les patients à leur domicile prématurément du point de vue clinique. Ainsi, dans tous les pays où la T2A a été introduite, on observe une réduction significative des durées de séjours. Néanmoins, de nombreux pays cherchant à réduire les durées de séjour à l'hôpital, avec ou sans T2A, la question qui se pose alors est de savoir lesquelles de ces réductions de durées sont « légitimes ». Le vrai défi est de pouvoir lier le paiement au patient pour qu'il reçoive les soins adéquats dans le cadre le plus adapté. Ce problème est bien défini par la théorie économique relative aux conséquences d'un système de T2A : lorsque le contenu du service associé à l'objet du paiement est mal défini (les soins fournis dans un séjour de GHM) le producteur tend à diminuer les services fournis dans le cadre du forfait et à transférer la responsabilité (et le coût) sur les autres (*cost shifting*).

Ainsi, aux États-Unis, la baisse de la durée de séjour à l'hôpital s'est accompagnée d'un accroissement démesuré des soins de suite et des soins à domicile (Newhouse, 2003). Il est donc important d'assurer un système de financement cohérent entre les différents modes de prises en

charge pour un même épisode des soins pour éviter ce phénomène de vases communicants (Cash *et al.*, 2003).

Les systèmes de santé 'intégrés', comme en Angleterre (NHS) où les généralistes assument la responsabilité totale, y compris financière, de la coordination de tous les soins fournis à leurs patients, quel que soit le cadre (à l'hôpital, dans les centres de soins et à domicile...), sont mieux armés pour atteindre ce défi. En Angleterre, le Ministre de la santé essaie également de « décomposer » la trajectoire des patients en séparant les parts relatives au diagnostic, à l'admission en court séjour, à la réhabilitation, etc. et en ajoutant des « sous-GHM » pour chacun de ces segments de l'épisode (Street *et al.*, 2007). En revanche, en France comme aux États-Unis, les médecins n'assurent pas la responsabilité de la coordination globale des soins et le financement des dépenses hospitalières est indépendant de celui des soins de ville.

Par ailleurs, dans ce contexte, la différence de champ dans la définition du contenu des soins affectés aux séjours entre les secteurs public et privé en France est un problème, car cela peut inciter les établissements privés à externaliser au maximum les examens et consultations diagnostiques ou pré-opératoires qui sont financés séparément sur l'enveloppe ambulatoire.

Une deuxième stratégie de réduction de la qualité des soins peut consister pour les établissements de santé à identifier à l'avance (avant l'admission) les patients pour lesquels les coûts générés par le séjour seront bien pris en compte au sein du tarif GHM et éventuellement décourager l'admission des autres patients moins « rentables » (phénomènes d'aversion au risque et de sélection des patients). Cela résulte de la variation des prises en charge spécifiques des patients au sein d'un même GHM et de donc l'hétérogénéité des coûts réels derrière un même tarif GHM. Les établissements peuvent aussi décider, en amont, d'éviter de se spécialiser dans les pathologies où la variation de *case-mix*³ est importante : cela a été bien démontré aux États-Unis où une grande partie des établissements ont décidé de fermer leurs services d'urgence entre 1983 et 1986 (Scanlon, 2006). Ce type de comportement est peu probable en France pour les établissements qui doivent satisfaire à une mission de service public et ceux qui ont une obligation de fournir les soins d'urgence à tous les patients (dont certains sont privés). Par contre, les autres établissements privés ont, en théorie, de fortes incitations financières à cibler les patients les moins graves ou pour lesquels le traitement est bien standardisé.

Troisièmement, les expériences étrangères montrent que la T2A entraîne un comportement de codage opportuniste (*upcoding*), au moins à court terme. Cela peut inciter également les établissements à fournir des soins ou des prestations inutiles afin de classer certains patients dans des GHM plus rémunérateurs. Ces codages fallacieux peuvent être repérés et corrigés à moyen ou long terme grâce à des contrôles efficaces.

³ Le *case-mix* désigne l'éventail des cas pris en charge et les niveaux de gravité des séjours associés.

3.2. Maîtrise des dépenses

La T2A incite à augmenter l'activité hospitalière. Ceci peut être souhaitable dans certains cas, notamment dans les systèmes où les délais d'attente pour la chirurgie programmée sont un réel problème, comme en Angleterre ou au Danemark. Mais ceci ne doit pas se faire de manière incontrôlée. Ce n'est pas non plus valable forcément pour l'activité médicale. De plus, la T2A génère un risque d'induction de la demande par les établissements, notamment pour les hospitalisations de jour et certaines chirurgies légères (Guterman, 2006). D'autre part, on sait que l'introduction d'un financement à l'activité à l'hôpital peut modifier l'équilibre des soins entre secteur hospitalier et secteur ambulatoire et ainsi avoir des répercussions sur les coûts médicaux en médecine de ville (Cash *et al.*, 2003).

Sans une planification rigoureuse et efficace de la capacité d'accueil et de prise en charge des patients, la T2A peut donc à la fois affaiblir la Maîtrise des dépenses hospitalières et distordre le financement des soins en fonction des besoins.

La capacité des financeurs (correspondant plutôt à celle des acheteurs dans les autres pays) à contrôler la demande de soins hospitaliers devient capitale avec la T2A. Lorsqu'il n'y a pas de seuil d'activité (défini *a posteriori*) imposé aux établissements, ceux-ci peuvent augmenter le volume des soins les plus profitables, éventuellement au détriment de la collectivité. En France, le contrôle des dépenses hospitalières globales est assuré par un mécanisme de baisse de tarifs en fonction de l'augmentation de l'activité globale. Il n'existe pas de seuils d'activité individualisés par établissement et une augmentation de l'activité globale ne reflète pas forcément le même gain d'efficacité pour tous les établissements. Ainsi, à enveloppe globale constante, certains établissements vertueux peuvent être pénalisés lorsque d'autres établissements, pratiquant le cas échéant la sélection de patients et l'induction de la demande, voient leurs recettes augmenter sans pour autant être plus efficaces du point de vue de la collectivité.

4. Les défis pour la mise en place d'une T2A

La qualité des données servant à la construction à et à l'application de la T2A, ainsi que les choix de classification et de méthode effectués, sont essentiels à son bon fonctionnement. Les deux éléments essentiels dans ce contexte sont les règles de la construction des GHM et le mode de calcul des coûts (sources de données employées, méthode de calcul, définition et prise en compte des séjours extrêmes). Une autre source de différence entre systèmes peut provenir du lien opéré entre les coûts et les tarifs. Nous dressons, ici, un bref récapitulatif des différentes options possibles et de leurs conséquences en termes d'incitation à l'efficacité.

4.1. La construction des GHM

Une définition homogène des prestations fournies est une condition primordiale de la T2A. Chaque pays utilise une classification différente. En France, il s'agit de la classification en groupes homogènes de malades ou GHM qui a été construite sur le modèle des *Diagnosis Related Groups* (DRG) américains. Le classement des séjours en GHM est effectué à l'aide d'une arborescence sur la base des diagnostics et actes techniques pratiqués au cours du séjour. Les actes sont considérés comme classants lorsque leur réalisation est décisive dans le coût d'un séjour.

À cet égard, les classifications utilisées aux États-Unis (Medicare DRGs) et en France sont très proches, avec quelques différences de classement pour certains actes. Cependant, Delattre *et al.* (2001) ont montré que les conséquences de ces différences de classement peuvent être non négligeables sur les stratégies de classements des patients, sur la mise en œuvre d'actes innovants et donc sur les tarifs de remboursement.

Par ailleurs, il est bien connu que les GHM sont plus adaptés à la description de patients requérant des soins aigus pour lesquels les coûts sont bien déterminés par le diagnostic ou la procédure de soins employée alors que ce type de nomenclature est moins fiable dans le cas de soins « médicaux », plus génériques et pour lesquels il est plus difficile de créer des groupes réellement homogènes (Street *et al.*, 2007). Dans la plupart des pays, les nomenclatures évoluent très régulièrement, dans le sens d'un accroissement du nombre de groupes tarifaires pour affiner les descriptions cliniques et mieux ajuster les tarifs correspondants.

Cela étant, chercher à réduire l'hétérogénéité des groupes en affinant de plus en plus les GHM et en multipliant ainsi le nombre de groupes et les tarifs correspondants va aussi à l'encontre du but initial poursuivi : celui de forfaitiser le coût des séjours selon des prises en charge homogènes et standardisées volontairement limitées en nombre. Ce principe initial répondait à l'objectif de changer les comportements de production hospitalière et d'optimiser/améliorer la pratique médicale.

4.2. Le calcul des coûts

Il faut pouvoir identifier correctement les coûts liés à un séjour hospitalier. Il se peut que la différence de coût observé entre les établissements ne reflète pas simplement les différences de coût réel mais également les différences dans les modes de comptabilisations des coûts des séjours (Sloan, 2000).

À ce titre, il est important de s'assurer que l'on puisse identifier les différences de coûts qui reflètent les différences de « pratique clinique » qui sont justifiées. Il est également important de pouvoir distinguer les coûts variables, générés par les soins directs reçus par le patient et que les établissements peuvent maîtriser ou ajuster, des coûts fixes sur lesquels ils ont peu de marge de manœuvre.

Par ailleurs la méthode utilisée pour mesurer les coûts a également un impact important sur les résultats. Concernant la performance des hôpitaux, il a été démontré qu'une partie des variations de

coûts observées provenait de différences de méthodologie employée pour calculer ces coûts (Tan *et al.*, 2009 ; Drummond *et al.*, 2005).

Les deux éléments clés de la méthodologie sont les suivants :

- le niveau de précision avec lequel les coûts sont identifiés, soit à un niveau macro (*gross costing*) soit à un niveau micro (*micro-costing*) ;
- le mode d'allocation des coûts aux différentes prestations, qui peut être réalisé selon une démarche descendante (*top-down*) ou ascendante (*bottom-up*).

Dans un modèle macro (*gross costing*), les coûts sont calculés à partir des données comptables des hôpitaux à un niveau agrégé (i.e. par journée) alors que dans un modèle de *micro-costing* tous les déterminants du coût sont mesurés au niveau le plus détaillé possible.

Dans une démarche descendante (*top-down*), les coûts sont distribués d'abord sur les services et ensuite sur les séjours à l'aide d'une clé de répartition, conduisant *in fine* à un coût moyen par patient. La pertinence de cette méthode dépend de l'exactitude des données financières et notamment de leur actualisation. À l'inverse, dans une démarche ascendante (*bottom-up*), les coûts sont calculés en identifiant les ressources réelles utilisées pour chaque patient conduisant à un coût unitaire spécifique patient. Ceci nécessite un système informatique comptable très détaillé pour rendre compte de tous les aspects des soins fournis pendant un séjour.

Le modèle de *micro-costing* est considéré comme supérieur pour mesurer les composantes de coût qui sont censées varier de manière significative entre les patients et pour les services qui présentent une part importante de coût de personnel (Swindle *et al.*, 1999 ; Wordsworth *et al.* 2005).

Mais, en pratique, des méthodes intermédiaires, combinant ces deux approches, sont souvent employées en exploitant au maximum l'information recueillie sur les patients pour développer des clés de répartition qui sont ensuite appliquées aux données administratives.

En France, les informations de coûts ayant permis la construction de la T2A proviennent uniquement des établissements publics (entre 45 et 50) qui ont participé aux différentes versions successives de l'étude nationale des coûts (ENC). Dans l'ENC, le recueil des coûts s'opère essentiellement selon une méthode descendante. L'avantage majeur de cette approche est d'éviter un investissement trop lourd en systèmes d'information comptable. En revanche, on sait qu'une telle méthode possède certaines limites (bien analysées dans la littérature) car les coûts y sont distribués en fonction d'une clé de répartition prédéterminée (*service weights*). On utilise en général des études auxiliaires pour déterminer le poids d'un GHM relativement à un autre GHM de référence (en termes de soins infirmiers consommés par exemple), mais cela crée souvent des problèmes de fiabilité (Phelan *et al.*, 1998). De plus certains coûts sont distribués dans l'ENC en fonction des durées de séjours hospitaliers ; or, cela ne tient pas compte des différences dans la gravité des cas traités ni dans l'intensité des soins médicaux pour un GHM spécifique.

De manière générale, l'expérience des autres pays montre que recueillir les données financières des établissements sans une évaluation détaillée des procédures utilisées peut conduire à des informations défectueuses et, *in fine*, à une échelle de coûts biaisée (Phelan *et al.*, 1998).

Dans le contexte de la T2A, la mesure des coûts constitue évidemment un enjeu important pour tous les pays. En France, cette question est capitale pour structurer le débat sur la convergence tarifaire public/privé. Or, paradoxalement, elle est en retard par rapport aux autres pays, d'une part en termes de collecte de données en routine du secteur public (pour lequel le nombre d'établissements observés est faible) et du secteur privé et d'autre part en termes d'analyse comparative des coûts entre établissements d'un même secteur et entre secteurs.

Aucun établissement privé n'a été observé dans les anciennes études nationales des coûts MCO. Les tarifs T2A appliqués en secteur privé ex-OQN (Objectif quantifié national) jusqu'à 2009 correspondaient aux moyennes des tarifs antérieurement facturés aux régimes d'Assurance maladie et non pas à des coûts réels observés. Rappelons ici que les tarifs du secteur privé ne couvrent qu'une part du coût du séjour, les honoraires des professionnels libéraux s'y ajoutant alors qu'ils forment une part importante et surtout très variable du coût réel du séjour.

En 2005, le Haut Conseil pour l'Avenir de l'Assurance maladie déplorait l'opacité des mécanismes de formation des coûts dans les établissements de santé et le manque de connaissances sur la question (HCAAM, 2005). Malheureusement, quatre ans après, il n'y a guère de progrès dans ce domaine. Cependant, une étude nationale de coût commune public/privé (ENCC) vient d'être publiée (octobre 2008) et doit permettre de mieux appréhender les comparaisons de coûts entre établissements publics et privés.

Cette échelle a servi d'ailleurs à la construction des tarifs 2009 à la fois pour les secteurs publics et privés.

4.3. Fixation des prix des GHM : le passage des coûts aux tarifs

Le niveau de prix (ou le tarif) correspondant aux différents GHM est un élément clé du fonctionnement de la T2A. Même si le principe de base est toujours de lier un prix à un GHM, la manière dont celui-ci est établi varie d'un pays à l'autre. Certains pays utilisent les coûts moyens comme référence de prix, les autres ont des benchmarks plus ambitieux (comme 10 % des établissements les plus efficaces).

En France et en Angleterre, les tarifs reflètent les coûts moyens dans la plupart des cas. Cela a un impact inflationniste car les établissements les plus performants sont encouragés à s'approcher de la moyenne plutôt que d'améliorer leur efficacité individuelle ; les moins performants ont une cible plus facile à atteindre mais qui ne correspond pas à leur optimum.

Déterminer le niveau de tarif approprié pour chaque GHM est donc un défi majeur pour la T2A. Si les prix sont trop faibles, les établissements apparemment les plus coûteux ne seront pas assez dotés

pour couvrir leurs coûts de production, ce qui peut conduire à des déficits hospitaliers, à des réductions des services et à des problèmes d'accès. Mais s'ils sont trop élevés, les dépenses d'Assurance maladie vont augmenter au profit des établissements apparemment les moins coûteux qui peuvent, en outre, afficher une activité accrue de manière artificielle.

Parallèlement, les politiques de paiement pourraient affecter négativement l'accès aux soins hospitaliers ou compromettre la qualité du service, en particulier le développement et la diffusion d'une nouvelle technologie.

Il n'est pas indispensable qu'il y ait un rapport direct entre les prix et les coûts réels. En fait, d'un point de vue théorique, les tarifs ne devraient refléter les coûts qu'à condition que la gamme existante des services de soins dans le marché soit idéale, ce qui est peu probable (Street et Maynard, 2007). D'ailleurs, l'objectif de la T2A est justement de changer la structure de production dans le marché pour améliorer l'efficacité. À ce titre, les prix peuvent être conçus et utilisés pour modifier le comportement des établissements en rendant certains types d'activités plus incitatives (financièrement) que d'autres. Par exemple, en Angleterre, les prix sont ajustés de manière à décourager l'hospitalisation classique là où la chirurgie ambulatoire est pertinente (voir la section suivante).

La littérature se fait l'écho d'une insatisfaction grandissante vis-à-vis des tarifs directement dérivés des coûts et de nombreux modèles mieux ciblés ont été proposés pour financer plus efficacement les coûts hospitaliers (Ellis et McGuire, 1990, Chalkley et Malcomson, 1998 ; Chalkley et Malcomson, 2002 ; Newhouse, 2002 ; Antioch *et al.*, 2007).

Du point de vue économique, un financeur (ou acheteur) public efficace paye le minimum nécessaire pour assurer un approvisionnement adéquat des soins. Un acheteur efficace est également conscient des différences en utilité ou en valeur des différents produits et services, et il intègre ces différences dans les prix qu'il paye aussi bien que dans la décision d'acheter ou non un service. La difficulté dans le secteur de santé est d'une part de spécifier et vérifier le contenu des produits (services de soins) sur lequel porte le paiement (ici le GHM) et d'autre part d'établir l'utilité relative des différents produits (soins) les uns par rapport aux autres.

Dans les pays où la T2A est en place depuis un certain temps, comme aux États-Unis et en Australie, on s'interroge sur la manière d'ajuster les prix afin de mieux refléter les différences de qualité des soins fournis. Ceci n'est pas simple et nécessite de développer des indicateurs spécifiques au traitement étudié et mesurés par patient. Néanmoins, on a observé que la plupart des facteurs déterminant la qualité sont liés au processus de soin et il est possible d'améliorer ce processus par des incitations financières. Ainsi, aux États-Unis, Medicare et Medicaid encouragent la participation aux programmes d'amélioration de la qualité en réduisant les tarifs des DRG (de -0,4 %) des établissements qui ne produisent pas les indicateurs de qualité demandés. De plus, Medicare a annoncé récemment qu'il ne paierait plus pour les séjours imputables à une non-qualité des soins fournis tels que certaines infections nosocomiales (AAWC, 2007).

En s'éloignant d'une tarification basée simplement sur les coûts (*cost based pricing*), il est possible de fournir de meilleures incitations pour entreprendre certains types d'activité tels que : donner la priorité aux traitements qui ont un plus grand impact sur l'état de santé, ou encore octroyer le financement dans un but d'amélioration de la qualité.

5. Enseignements des expériences étrangères de T2A

Nous présentons ici quelques exemples de la manière dont différents pays ont instauré leur système de T2A et établi leurs tarifs. À partir des expériences de trois pays, États-Unis, Angleterre et Australie, nous examinerons en particulier, comment ces systèmes sont confrontés aux défis principaux : la définition des GHM et leur coût et la détermination des tarifs.

5.1. États-Unis

Contexte du système de soins

La tarification à l'activité a été introduite la première fois aux États-Unis en 1983 pour payer les séjours des patients âgés couverts par le programme Medicare. L'objectif principal était de contrôler l'augmentation des dépenses hospitalières qui avaient presque doublé entre 1970 et 1982 lorsque les hôpitaux étaient remboursés sur la base de leurs coûts réels. Sous ce système, un dollar dépensé était égal à un dollar remboursé. Aujourd'hui, tous patients confondus, les hôpitaux américains reçoivent un tiers de leur revenu par Medicare. Un autre tiers vient des assureurs privés qui utilisent les mêmes groupements en DRG (équivalents de nos GHM) mais négocient leurs tarifs annuellement avec chaque établissement. Le reste vient du programme Medicaid (~20 %) et des paiements directs des patients qui n'ont pas d'assurance. Ainsi, un hôpital peut être financé à des niveaux de prix différents pour une même prestation car celle-ci aura été négociée différemment avec plusieurs acheteurs.

Malgré les variations et les marges de manœuvre importantes laissées aux acheteurs, les pratiques de tarification de Medicare constituent la référence en la matière. Pour mieux comprendre comment cela marche, nous rappelons ci-dessous les principaux éléments de la politique de financement de Medicare.

L'unité de paiement : DRG

Medicare utilise les groupements de diagnostics nommés DRG pour payer les séjours en soins médicaux et chirurgicaux. Les tarifs DRG couvrent tous les coûts directs attribuables aux séjours d'un patient. Les coûts directs liés à l'enseignement et à la recherche, aux soins ambulatoires, à la réhabilitation, aux soins du cancer et à la psychiatrie ne rentrent pas dans le champ du paiement par DRG. Le nombre de DRG est passé de 470 en 1983 à 557 en 2006. En octobre 2007, Medicare a introduit un nouveau système plus spécifique en ajustant mieux par la sévérité des cas traités (MS-

DRGs) afin de mieux cibler les paiements (EHM, 2007). Ainsi, il divise un unique DRG avec complications en deux MS-DRG : avec complications majeures et avec complications courantes. Aujourd'hui il y a 999 MS-DRGs. Notons également que la classification '*All-patients – refined DRGs*' utilisée pour les patients non-couverts par Medicare (<65 ans) porte sur 1 100 catégories et prend en compte non seulement la sévérité de la maladie, mais également le risque de mortalité mesuré sur quatre niveaux (mineur, modéré, majeur, extrême).

Le calcul des coûts : DRG weights

Medicare sépare les calculs de coût de la grille de tarifs établie. Pour chaque DRG, un coût moyen est calculé en prenant la moyenne des ressources consommées par tous les patients Medicare inclus dans ce DRG. Ces coûts sont standardisés pour éliminer les différences de coûts de structure et de personnel dans différentes régions. Ceci fournit un « tarif de base » qui est ensuite pondéré par deux indices et ajusté par les coûts spécifiques supportés par les différents types d'établissements. Le coût direct de l'enseignement médical est payé séparément. Par ailleurs, les déficits hospitaliers dus au non-paiement de factures ne sont pas inclus dans ces calculs de coûts.

Medicare effectue des paiements en sus pour les cas qui sont extrêmement lourds et chers (*outliers*) en sélectionnant les séjours dont le coût s'écarte de la moyenne de plus de trois fois la valeur de l'écart-type.

La définition des tarifs des DRG

Les tarifs des DRG sont ajustés pour tenir compte de l'hétérogénéité des offreurs de soins (Hsiao *et al.* 1986). Quatre facteurs sont considérés comme déterminants des coûts des services fournis par l'hôpital :

1. Les facteurs exogènes du marché (salaires et biens d'équipement) : les niveaux des salaires du personnel hospitalier étant très différents entre les régions ou les zones géographiques, Medicare pondère les tarifs de base par un indice salarial qui distingue notamment les zones urbaines et rurales. Les zones où les salaires sont plus élevés obtiennent un tarif plus élevé. Un facteur d'ajustement géographique est également calculé d'une manière identique en évaluant le coût relatif des autres biens et services de la zone (Huang et Frank, 2006).
2. Le coût indirect de la formation médicale (enseignement) : les établissements qui font de l'enseignement sont supposés avoir des coûts plus élevés que les autres établissements en raison des tests et examens (examens complémentaires, certaines procédures cliniques utilisées dans une optique pédagogique) et par le fait qu'ils prennent en charge des patients en moyenne plus lourds. Ainsi, le tarif des DRG pour ces établissements est augmenté d'un certain pourcentage déterminé par le nombre moyen d'internes par lit.
3. Le paiement pour la prise en charge de populations spécifiques : les établissements qui consacrent une fraction importante de leur activité à la prise en charge des populations à faible revenu reçoivent des paiements spécifiques pour cette activité.

4. En plus de ces trois facteurs, les tarifs DRG sont ajustés en prenant en compte la catégorie de l'établissement. Il existe des règles spécifiques si l'hôpital est considéré comme l'unique hôpital d'un bassin de population, s'il s'agit d'un établissement rural dépendant de Medicare ou d'un hôpital régional de référence (Scanlon, 2006). Un hôpital est désigné comme « unique hôpital d'un bassin de population » s'il satisfait aux trois conditions suivantes : 1) il n'y a pas d'autre établissement dans un rayon de 35 miles ; 2) il est la seule source de soins aigus dans cette zone géographique ; 3) il est désigné par le Ministère comme « hôpital d'accès critique » (*critical access hospital*). Un établissement rural dépendant de Medicare est un hôpital qui perçoit au moins 60 % de ses revenus par Medicare. Un centre régional de référence est un établissement qui accueille un nombre significatif de patients des autres établissements de la région.

Ajustement des DRG

Afin d'éviter des incitations néfastes qui pourraient induire une sélection des patients ou des choix de traitements *a priori*, les paiements des DRG sont ajustés au coût prévu pour le séjour d'un patient. Mais, dès la mise en place du système de T2A, on a constaté que la classification des DRG n'expliquait pas toute la variation de coût entre les différents séjours : en fait, les DRG n'expliquaient qu'environ 30 % de cette variation (chiffre équivalent à ce qui a été calculé pour la France lors de l'introduction des GHM).

En 1995, la Commission d'évaluation du paiement à l'activité (ProPAC) a recommandé de modifier les DRG pour mieux distinguer dans les paiements les séjours/patients en tenant compte des diagnostics secondaires spécifiques qui affectent souvent le coût de soins, ainsi que les différences de pratique médicale entre les établissements.

Malgré des ajustements successifs, une étude réalisée pour MedPAC en 2005 a eu des conclusions similaires. Cette étude démontre des variations importantes en termes de rentabilité pour un même DRG et entre les DRG. À contexte clinique similaire (diagnostics principaux notamment), les DRG concernant la chirurgie seraient plus profitables que des traitements médicaux, notamment pour des soins orthopédiques et de cardiologie, cette variation de coût étant systématiquement liée au diagnostic secondaire du patient. Cette étude a conduit à une recommandation de MedPAC visant à améliorer l'ajustement du *case-mix* à travers les hôpitaux, et à ajuster leurs prix en affinant encore les DRG pour mieux capturer les différences de sévérité des cas traités (Scanlon, 2006).

Limites du raffinement des DRG

La discussion actuelle aux États-Unis portant sur le rôle des hôpitaux et l'impact potentiel de la poursuite du raffinement des DRG sur certains hôpitaux est instructive. D'abord, les experts ont noté que le raffinement introduisait plus de complexité dans la classification des DRG, nécessitant une base d'information très détaillée et sophistiquée, avec un risque accru de « mal payer » lorsque ces informations ne sont pas correctement collectées.

Mais on a surtout mis en avant le fait que raffiner davantage les DRG risquait d'avoir des effets pervers sur la dotation des établissements : les analyses de ProPAC ont ainsi révélé que ce raffinement allait conduire à payer plus les grands établissements d'enseignement et moins les petits établissements ruraux, alors même que ces derniers ont déjà du mal à survivre dans le système actuel.

Notons finalement que la théorie économique prévoit que le pouvoir du financeur diminue lorsque le nombre de catégories de paiement augmente dans un système de paiement prospectif (Newhouse, 2003 ; Cash *et al.*, 2003). Plus le nombre de tarifs augmente en différenciant non seulement les situations cliniques mais aussi les choix thérapeutiques, plus le financeur (acheteur) perd en pouvoir de négociation.

5.2. Angleterre

Contexte du système de soins

L'Angleterre dispose d'un système national de santé (NHS) où les services de santé sont financés par les impôts collectés au niveau national pour toute la population. L'accès aux soins est gratuit et assuré pour tout le monde. Mais l'accès aux soins de spécialistes en ville et à l'hôpital est contrôlé par les médecins généralistes (appelés GP) qui assurent ainsi un rôle de « *gatekeeper* ». Le financement des soins primaires (médecins généralistes) et secondaires (spécialistes et hôpitaux) est sous la responsabilité des PCT (*Primary care trusts*) qui gèrent 80 % des budgets du NHS. Les PCT reçoivent leur budget du Ministère de la santé et sont responsables de l'achat de l'ensemble des soins médicaux pour la population qu'ils couvrent.

Les longues durées d'attente pour l'accès à la chirurgie programmée ont été le problème majeur du système de santé en Angleterre depuis les années quatre-vingt. De nombreuses réformes ont cherché à résoudre ce problème. La T2A, ou le « paiement aux résultats » comme on dit en Angleterre, fut introduite en 2004 pour augmenter la productivité hospitalière et accroître notamment l'activité programmée (Audit Commission, 2005). Les PCT paient pour l'activité hospitalière avec des prix/tarifs fixés au niveau national mais ils gardent une petite marge de manœuvre pour négocier les volumes d'activité avec chaque établissement. Avant la T2A, il y avait un certain nombre d'arrangements contractuels (plus ou moins compliqués) entre les PCT et les établissements définissant le volume global de l'activité pour une gamme de services et de soins définis.

L'unité de paiement : HRG

L'Angleterre utilise ses propres groupes homogènes de ressources (*Healthcare Resource Groups*, HRG) pour payer l'activité. La première version (3.5) comprenait 550 groupes homogènes ; la dernière version (HRGv4) qui vient d'être actualisée en 2009 compte plus de 1 400 groupes (The Information Center, 2009).

Le calcul des coûts

Le système comptable en Angleterre est assez similaire à celui de la France. Les coûts sont calculés à partir des budgets (*top down costing*) et il n'y a pas de données individuelles au niveau patient. En revanche, l'information sur les coûts hospitaliers provient de la totalité (100 %) des hôpitaux concernés par l'application de la T2A (300 *Hospital Trusts*⁴) alors qu'elle ne porte que sur un échantillon restreint d'établissements en France (seulement 3 % des établissements concernés par la T2A).

La définition des tarifs

La T2A a été introduite de manière progressive. Elle s'appliquait au début (2003/2004) aux 15 HRG concernant la chirurgie programmée (ils sont passés à 30 en 2004/2005). Les tarifs ont été établis de manière à augmenter le niveau d'activité au-delà d'un niveau pré-négocié (*Service level agreement*). Il était prévu que le tarif serait revu à la baisse pour les établissements qui n'ont pas réussi à fournir ce niveau minimum d'activité (Farrar *et al.*, 2007).

Les tarifs correspondent *a priori* aux coûts moyens d'une procédure de soins calculés sur la population totale des hôpitaux. Toutefois, cette règle est modifiée dans certains cas en suivant des objectifs spécifiques (Street *et al.*, 2007) :

- Les paiements aux établissements pour un HRG donné sont modulables selon que le séjour était programmé ou pas en prenant en compte les surcoûts pour les séjours non-programmés.
- Quand la chirurgie de jour est pertinente, le tarif retenu est la moyenne pour l'activité de chirurgie de jour, donnant de ce fait aux établissements de véritables motifs pour développer une telle activité, ce qui représente un enjeu politique historique en Angleterre. D'ailleurs, il n'y a pas de contrainte sur les volumes d'activité chirurgicale puisque le gouvernement vise à réduire les durées d'attente.
- Cependant, en ce qui concerne l'activité de soins des services d'urgences, le plein tarif est payé pour un volume d'activité égal à celui de l'année passée plus 3 % maximum. Une fois ce volume d'activité dépassé, les tarifs déclinent de 50 %. L'objectif de ce mécanisme est de contrôler la demande de soins aux urgences hospitalières et de développer les équipements alternatifs de soins de premier recours dans la communauté.
- Les coûts de référence, qui sont calculés en moyenne nationale, sont ajustés avec un indice du marché local (*Market forces factor*) pour moduler un tarif local en fonction de l'environnement économique de l'établissement.
- Pour certains HRGs, les tarifs sont ajustés en prenant en compte les recommandations de l'Institut National d'Excellence Clinique (NICE) sur les traitements et les médicaments qui sont coûts-efficaces.

⁴ Les *Hospital trusts*, créés en 1991 pour introduire les règles du marché en Angleterre, sont des entités juridiques qui gèrent souvent plusieurs établissements.

5.3. Australie (province de Victoria)

Contexte du système de soins

Tous les Australiens ont accès aux soins à travers le plan Medicare introduit en 1984 et financé par les impôts. La responsabilité financière et administrative de Medicare est partagée entre le gouvernement fédéral et les six gouvernements régionaux (états). Même si le financement provient à la fois du gouvernement fédéral et des états, les hôpitaux publics sont payés exclusivement par le gouvernement régional, qui assure seul la régulation. Victoria est un de plus grand état d'Australie couvrant plus de 25 % de la population (environ 4,5 millions de personnes). Il est le premier état à avoir introduit la T2A en 1993 pour remplacer le remboursement sur la base des coûts historiques des établissements. Les durées d'attente en chirurgie constituaient alors un problème majeur en Australie, le nombre des patients en attente ayant souvent été utilisé comme un élément de négociation par les hôpitaux pour demander des financements supplémentaires.

L'unité de paiement : AN-DRG

L'Australie a développé une version plus raffinée des *Yale DRGs* utilisés par le Medicare américain. Cette nomenclature a été créée dans les années quatre-vingt-dix et adoptée dans la plupart des états australiens. La classification des séjours intègre une hiérarchie des prestations et des situations médicales. Les séjours sont distribués dans 23 classes médicales et dans 667 AN-DRGs. L'état de Victoria a modifié de manière progressive les AN-DRG, jusqu'à créer ses propres VIC-DRGs. Ceux-ci ajustent notamment les paiements des séjours de moins de 24 heures en fonction de la difficulté technique du traitement. Il y a actuellement 760 VIC-DRGs, mais il existe un mécanisme de révision annuel encadré par un panel de cliniciens.

Le calcul des coûts

Les coûts des séjours sont calculés à partir des données aux niveaux patients compilés par une méthode ascendante (VGHI, 2009). Ce type de modèle agrège toute l'information disponible au niveau des patients par spécialité médicale dans des centres des coûts dédiés à ce travail (*cost centers*). Ce type de modélisation est censé améliorer l'exactitude de la mesure des coûts des patients comparativement aux méthodes dans lesquelles les coûts sont attribués aux patients en fonction des durées de séjours seules.

La définition des tarifs :

Avant l'introduction de la T2A les hôpitaux recevaient des budgets globaux en fonction de leur activité propre et le nombre de patients dans leur liste d'attente. Ce mécanisme a été critiqué car il incitait les établissements à « garder » les patients dans leur liste d'attente plutôt que le contraire. Un des objectifs principaux de l'introduction de la T2A a donc été de réduire le nombre de patients dans les listes d'attentes ; la politique de tarification mise en place a été déterminée en fonction.

Notons d'abord que des tarifs uniques par DRG sont appliqués aux « hôpitaux similaires » en prenant en compte leur *case-mix* et utilisés pour négocier des contrats prix-volume. Les prix sont également ajustés par un grand nombre de paramètres, y compris les conditions sociodémographiques du bassin de clientèle d'un établissement (Brook, 2007). Chaque hôpital reçoit en général un prix fixe par DRG jusqu'à un certain niveau d'activité cible déterminé en fonction de son activité historique. Mais comme le gouvernement voulait inciter les établissements à dépasser leur niveau d'activité cible, des mécanismes de paiements additionnels ont été mis en place :

- D'abord, l'objectif étant d'orienter l'activité pour prendre en charge les patients qui ont une urgence clinique et/ou ceux qui sont en attente de soins depuis longtemps, le Ministère a développé – en collaboration avec les cliniciens – un système de classification des patients en fonction du degré d'urgence de leur besoin clinique (urgent, semi-urgent ou non-urgent).
- Afin de maîtriser les dépenses hospitalières totales, le Ministère a plafonné les paiements pour l'activité additionnelle à une enveloppe fixe (P) qui prévoyait, en 1993/94, un financement additionnel pour un supplément d'activité de 7 % au total. Les fonds de cette enveloppe sont distribués à chaque établissement *au prorata* de sa contribution à l'activité additionnelle totale. Les tarifs payés pour l'activité additionnelle sont donc fluctuants : ils sont calculés en divisant l'enveloppe globale à répartir (P) par le nombre de séjours additionnels réalisés.
- Finalement, le gouvernement voulait s'assurer que les établissements donnent bien la priorité aux patients qui figuraient dans leurs listes d'attente. Pour cela, l'accès aux financements supplémentaires (P) a été conditionné à la réalisation d'objectifs spécifiques de prise en charge des patients sur liste d'attente. Par exemple, la première année de mise en place de ce système, les établissements ne pouvaient pas bénéficier des paiements pour activité additionnelle s'ils n'avaient pas soigné tous les patients identifiés comme « urgents ».

Cette politique de tarification s'est avérée plutôt efficace (Street *et al.*, 2007), avec des réductions drastiques dans le nombre de patients « urgents » attendant plus de 30 jours.

Ceci fournit un exemple de politique active de tarification et montre comment les régulateurs peuvent influencer les incitations fournies par une T2A en fonction des objectifs et contraintes propres à leur système.

6. Conclusion

Le principe de base de la tarification à l'activité – payer les établissements en fonction de leur activité mesurée par Groupe Homogène de Malades (GHM) – a un sens économique. Toutefois, la littérature et les expériences des autres pays présentés dans ce papier montrent que ce mécanisme de paiement présente un certain nombre de risques et nécessite des ajustements réguliers et soigneux pour obtenir les bénéfices attendus d'un tel système.

Premièrement, la classification de l'activité hospitalière dans des groupes homogènes est un véritable défi de la T2A. Cette classification requiert de fréquents ajustements et un arbitrage entre son niveau de précision et son applicabilité en termes de grille tarifaire. Du point de vue technique, les GHM doivent se fonder sur des groupes qui ont une cohérence aussi bien économique que clinique (Busse *et al.*, 2006). Ceci nécessite des décisions minutieuses et transparentes concernant la construction des algorithmes qui classent les patients dans les différents groupes.

Deuxièmement, en tant que mode de financement, la T2A n'a aucune vocation à assurer une couverture optimale des besoins ni à améliorer la qualité des soins. Au contraire, il est nécessaire de prévoir des mécanismes régulateurs complémentaires pour garantir la qualité des soins et un accès équitable. Cela est particulièrement important dans le contexte de soins en France où le secteur privé à but lucratif joue un rôle conséquent dans la prise en charge de certaines pathologies, alors que la régulation des prix n'y est que partielle en raison des dépassements d'honoraires non couverts par l'assurance obligatoire.

Troisièmement, dans une perspective de maîtrise des dépenses de santé, il est important de suivre de près l'évolution de l'activité dans les différents modes de prise en charge hospitalière, ainsi qu'en médecine de ville. Par nature, la T2A incite les établissements à augmenter leur activité ; l'expérience des autres pays montre que l'introduction de la T2A a accru la propension des établissements à induire la demande de soins pour développer l'activité de certains de leurs services profitables (notamment en chirurgie légère et en chirurgie ambulatoire). Cette modification dans le mode de financement et dans les usages hospitaliers peut également occasionner des transferts d'activité massifs vers les soins de suite ou vers domicile et même modifier l'équilibre ville/hôpital.

Cela étant, le mécanisme de régulation prix/volume actuellement en vigueur en France (avec les enveloppes ONDAM) ne rend pas compte de la réalité des évolutions d'activité médicale. Compte-tenu de sa relative opacité et de sa nature très macro-économique, il risque de ne pas assez « récompenser » les efforts réalisés au niveau de chaque établissement considéré individuellement, ce qui, au total, pourrait être contre-productif.

Quatrièmement, le principe de payer un prix fixe qui soit directement indexé sur les coûts moyens observés et commun à tous les types d'établissements est de plus en plus contesté. Ce mode de fixation des tarifs n'est adapté ni pour améliorer l'efficacité globale du marché hospitalier, ni pour inciter chaque établissement à améliorer sa propre efficacité. La littérature évoque plusieurs méthodes alternatives qui permettent de mieux cibler le financement des activités médicales en privilégiant, notamment, les soins qui apportent le meilleur rapport coûts-bénéfices. Un tel système suppose que les acheteurs/financeurs du système soient plus proactifs dans la négociation des prix et qu'ils mènent des politiques volontaristes pour structurer la gamme de soins et de services offerts en fonction des besoins de populations spécifiques et du niveau de qualité de ces soins. En France, cela nécessiterait que l'Assurance maladie (et ses représentations locales) adopte une telle position pour affirmer davantage son rôle d'« acheteur » de soins à part entière et non uniquement de « payeur ».

La performance de tout système de paiement est appréciée en définitive par sa capacité à fournir des incitations appropriées aux producteurs de soins (les établissements de santé et les médecins), qui sont en ligne avec les objectifs sociaux. La T2A peut engendrer des comportements opportunistes de la part des établissements et pourrait affecter négativement l'accès aux soins hospitaliers ou en compromettre la qualité. Une politique efficace, s'appuyant sur un recueil d'informations détaillées permettant de comprendre les différences de pratiques médicales et de suivre les changements dans les comportements des différents acteurs, est essentielle pour assurer le succès d'un système de tarification à l'activité.

Remerciements

Cette étude a partiellement bénéficié d'un financement de la Fédération Hospitalière de France (FHF). Les auteurs tiennent à remercier Laure Com-Ruelle pour son expertise et sa collaboration tout au long de cette étude, ainsi que Julien Mousquès, Chantal Cases et Thierry Debrand pour les nombreux commentaires.

Bibliographie

AAWC (2007), Changes to Medicare hospital payment 2007-2008, Association for the Advancement of Wound care, on line, <http://www.aawonline.org/medicare-payment.shtml>

Antioch K, Ellis R, Gillet S, Borovnicar D, Marshall R. (2007), "Risk adjustment policy options for case-mix funding: international lessons in financing reform", *European Journal of Health Economics*, vol. 8(3), pp. 195-212.

Audit Commission (2005), Early lessons from payment by results, National Report, Octobre 2005. Disponible : www.audit-commission.gov.uk

Brook C. (2007), *Casemix funding for acute hospital care in Victoria, Australia* <http://www.health.vic.gov.au/casemix/about.htm>, State Government of Victoria, The Department of Human Services.

Busse R., Schreyögg J. and Smith P. (2006), Hospital case payment systems in Europe, *Health Care Management Science*, **9**: 211-13.

Cash R., Grignon M. et Polton D. (2003), "L'expérience américaine et la réforme de la tarification hospitalière en France", Commentaire de l'article de J. Newhouse, *Economie publique* n°13, 3-15.

Chalkley M. and Malcomson J. M. (1998), 'Contracting for health services when patient demand does not reflect quality', *Journal of Health Economics*, 17(1): 1-19.

Chalkley M. and Malcomson J. M. (2000), 'Government purchasing of health services', in Culyer A. J. and Newhouse J. P. (eds), *Handbook of Health Economics*, Vol. 1A, Amsterdam: Elsevier Science, chapter 15, pp. 847-890.

Delattre E., Dormont B., McClellan M. and Milcent C. (2001), Systèmes de tarification et évolutions de la variabilité des coûts hospitaliers en France et aux États-Unis – Texte Intégral. *Document de travail DREES*, 2001.

Drummond M., Sculpher M. J., Torrance G. W., O'Brien B. J. and Stoddart, G. L. (2005), *Methods for the economic evaluation of health care programmes*, 3rd edn, Oxford University Press, New York.

EHM (2007), Executive Healthcare Management, MS-DRGs, What is at stake for hospitals and physicians?, <http://www.executivehm.com/currentissue/article.asp>, issue 226.

Ellis, R. P. and McGuire, T. G. (1990), 'Optimal payment systems for health services', *Journal of Health Economics*, 9(4), 375-96.

Farrar S., Sussex J., Yi D., Chalkley M., Scott T., Sutton M. and Yuen P. (2007), "National Evaluation of Payment by Results", *Health Economics Research Unit*, Report to the Department of Health.

Guterman S. (2006), "Specialty Hospitals: A Problem or a Symptom?", *Health Affairs*, Volume 25 (1): 95-105.

HCAAM (2005), Haut Conseil pour l'Avenir de l'Assurance maladie, Document de travail sur l'hospitalisation, novembre 2005.

Hsiao W., Sapolsky H., Dunn D. and Weiner S. (1986), 'Lessons of the New Jersey DRG Payment system', *Health Affairs*, Summer 1986, pp. 32-45.

Huang Z. and Frank B. (2006), *Calculating Hospital Specific DRG Payments*, Technical Brief, ResDAC Publication Number TN-004-02, November 2006. Research Data Assistance Center, University of Minnesota, Minneapolis, MN. <http://www.resdac.umn.edu>.

Newhouse, J.P. (2002), "Pricing the Priceless: A Health Care Conundrum", Cambridge, MA: MIT Press.

Newhouse J.P. (2003), "Reimbursing for Health Care Services", *Economie publique*, 13: 3-31.

Or Z., Renaud T. et Com-Ruelle L. (2009), 'Les sources de variation des coûts hospitaliers : réflexions sur une convergence tarifaire entre les secteurs public et privé', *Document de travail IRDES*, à paraître.

Phelan P.D., Tate R., Webster F. and Marshall R.P. (1998), "DRG cost weights – getting it right". *EMJA Casemix: moving forward*; 169: S36-S38.

Scanlon W.J. (2006), The Future Of Medicare Hospital Payment "Modest" proposals in light of Medicare's challenges. *Health Affairs – Vol. 25 (1)*: pp. 70-80.

Sloan F. (2000), 'Not for profit ownership and hospital behaviour', dans A. Culyer and J. Newhouse (eds), *Handbook of Health Economics*, Elsevier, Amsterdam, pp. 1141-1174.

Street A., Maynard A. (2007), "Activity based financing in England: the need for continual refinement of payment by results", *Health Economics, Policy and Law*, 2: 419-427.

Street A., Vitikainen K., Bjorvatn A., Hvenegaard A. (2007), "Introducing activity-based financing: a review of experience in Australia, Denmark, Norway and Sweden", *CHE Research Paper 30*, Centre for Health Economics, University of York.

Swindle R., Lukas C., Meyer D., Barnete P., Hendricks A. (1999), 'Cost analysis in the department of Veterans affairs: consensus and future directions', *Medical Care*, 37(4 supp), AS3-8.

Tan S.S., Rutten F.F.H, van Ineveld B.M, Redekp W.K and Hakkaart-van Roijen L. (2009) Comparing methodologies for the cost estimation of hospital services, *The European Journal of Health Economics*, 10: 39-45

The Information Center (2009), *Healthcare Resource Groups (HRGs)*

www.ic.nhs.uk/services/casemix/hrg4/hrg4-reference-cost-grouper-download

VGHI (2009), Victorian Government Health Information, *Casemix funding in Victoria*, en ligne:

www.health.vic.gov.au/casemix/definitions

Wordsworth S., Ludbrook A., Caskey F. and McLeod A. (2005), 'Collecting unit cost data in multicenter studies. Creating comparable methods', *European Journal of Health Economics*, 6(1), 38-44.