

Les indicateurs souvent utilisés pour mesurer et comparer la qualité des soins dispensés par les hôpitaux – volume d'activité, mortalité hospitalière, attractivité de l'établissement, durée moyenne de séjour, taux de césariennes et d'infections nosocomiales – font ici l'objet d'une analyse critique. Cette analyse de leurs conditions d'utilisation et de leur validité est issue des travaux d'un groupe d'experts et de scientifiques. Elle met en évidence le caractère pluridimensionnel des notions de qualité et de performance hospitalières, ainsi que la nécessité de les mettre en perspective avec les missions et l'environnement propres aux différents types d'établissements. La recherche de meilleurs indicateurs et de modes de comparaison adaptés passe alors par une forte implication des professionnels. La démarche statistique consisterait, dans ce cadre, à envisager un ensemble large et pluriel d'indicateurs pour des établissements « homogènes » en termes d'activité, de contraintes et d'environnement.

Pierre LOMBRIL

PU-PH en santé publique
CHU Nantes

Michel NAIDITCH

Médecin, maître de conférence Université Paris VII,
chercheur groupe IMAGE ENSP

Dominique BAUBEAU

Ministère de l'Emploi et de la solidarité
DREES

Philippe CUNEO

Ministère de l'Emploi et de la solidarité
DREES

Avec la collaboration de **Jean CARLET**
Président du Comité technique national
des Infections nosocomiales

Les éléments de la « performance » hospitalière les conditions d'une comparaison

Un ensemble de publications a récemment présenté des classements des établissements de santé en fonction de leur « qualité ». Ces démarches partent d'une idée simple : la qualité technique des soins est mesurable de façon directe ou indirecte, et il est possible de fabriquer un score à partir de plusieurs « notes » que l'on combine en leur affectant un certain coefficient. Ces démarches appellent du point de vue de l'analyse statistique des réflexions qui s'articulent en trois points :

– La qualité est une notion multidimensionnelle : selon la définition de l'Organisation mondiale de la Santé (OMS), l'évaluation des soins doit permettre de « garantir que chaque patient reçoive la combinaison d'actes diagnostiques et thérapeutiques qui lui assurera le meilleur résultat en termes de santé, conformément à l'état actuel de la science médicale, au meilleur coût pour un même résultat, au moindre risque iatrogène¹ et pour sa plus grande satisfaction en termes de procédures, de résultats et de contacts humains à l'intérieur du système de soins ». Les approches de la qualité doivent donc comporter

1. Iatrogène : qui est provoqué par le médecin ou par le traitement médical.



ces dimensions plurielles ; les approches partielles à partir d'un ou deux indicateurs posent des problèmes de validité.

– Par ailleurs, la construction d'échelles de pondération entre un ensemble d'indicateurs de nature différente apparaît problématique. Elle équivaut, en effet, à établir une hiérarchie de valeurs entre la satisfaction des patients, l'absence de complications, les résultats sur l'état de santé, la pertinence des actes ou le coût, hiérarchie qui reflète des jugements subjectifs non exprimés et non soumis à débat.

– Si le principe d'un classement de la « qualité globale » des hôpitaux est méthodologiquement contestable, il semble néanmoins intéressant et possible de chercher à déterminer un ensemble d'indicateurs dont les conditions d'utilisation à des fins de comparaison soient établies sur des bases scientifiques ; ce qui implique de comparer des structures qui partagent un type d'activités et de contraintes similaires et d'analyser séparément chacun des domaines qui la constituent.

C'est pourquoi il a paru nécessaire de faire le point sur les indicateurs les plus fréquemment cités, leur intérêt, leurs limites et leurs conditions

d'utilisation. Cette première synthèse s'appuie sur les travaux d'un groupe d'experts réuni à l'initiative de la DREES (encadré 1) ; elle sera complétée, au fur et à mesure de l'avancement de leurs travaux, par des développements sur les différentes dimensions de la performance hospitalière (encadré 2) et les moyens d'en rendre compte.

Quelle validité des indicateurs habituels de comparaison des établissements hospitaliers ?

En ce qui concerne les indicateurs censés donner des informations sur la qualité au sens donné par l'OMS, les plus souvent cités sont : le volume d'activité, le taux de mortalité, l'attractivité, la durée moyenne de séjour, le taux d'infections nosocomiales et le taux de césariennes.

■ Le volume d'activité

Il est souvent considéré comme un « indicateur de bon sens » indiscutable. Il pose néanmoins des problèmes importants :

– D'abord, s'agit-il d'actes tous nécessaires et appropriés ? En France comme à l'étranger, les enquêtes mettent en évidence « la variabilité des pratiques médicales » : pour une

même affection et dans un même contexte, l'utilisation des actes et techniques médicales varie de façon importante selon les équipes. En chirurgie osseuse² comme en chirurgie cardiaque³, le nombre d'interventions par habitant varie du simple au double en France selon les régions, ce qui signifie des variations entre établissements plus élevées et suggère l'existence de phénomènes de sur et de sous-utilisation de ces actes. Pour tenter de remédier à cette situation, tous les pays ont développé des références de bonne pratique clinique (RPC). En France, la mission confiée à l'Agence nationale d'accréditation et d'évaluation en santé (ANAES), le développement de recommandations à l'initiative de diverses fédérations ou sociétés savantes et la multiplication de « règles de pratique⁴ » au sein des hôpitaux témoignent du même souci de réduire les écarts de pratique et de promouvoir une « médecine basée sur les preuves scientifiques ». Le chiffre de l'activité ne permet donc pas de garantir à lui seul la qualité des prestations, dès lors qu'il n'a pas été fait la preuve que les actes réalisés constituaient les traitements les plus adaptés.

– L'existence d'une relation systématique entre le niveau d'activité et le résultat des soins n'a pas été démontrée scientifiquement. La majorité des études qui ont défini des seuils minimaux d'activité (appelés « masse critique ») n'ont en effet pas tenu compte des différences entre les populations traitées selon les établissements. De plus, les critères utilisés pour définir la qualité des soins délivrés, notamment les taux de mortalité, étaient souvent discutables. Des preuves convaincantes de l'effet d'un

2. Enquête GAIN Chirurgie, ENSM-CNAMTS 1995.

3. Activité des unités de chirurgie cardiaque en 1995, Étude DH-CNAMTS-IMAGE.

4. Encore appelées « guidelines ». Il s'agit de documents destinés à guider la démarche du médecin prescripteur.

E-1

Pour réaliser cette première synthèse, la DREES s'est appuyée sur un groupe de scientifiques et d'experts chargé de faire le point sur les méthodes et les indicateurs utilisables pour comparer les performances des établissements de santé. Ce groupe restreint d'experts permanents composé de personnalités nommées « intuitu personae » a été créé en mai 1999. Il est présidé par P. LOMBRAIL, Professeur de santé publique au CHU de Nantes, assisté de M. NAIDITCH, enseignant-chercheur en santé publique du groupe IMAGE (ENSP).

Ce groupe multidisciplinaire comprend :

- des économistes : A. Letourmy (CERMES) et S. Charvet-Protat (Union Hospitalière Privée),
- des spécialistes en santé publique et en information médicale : M.-J. Sourty-Le Guellec (CREDES), J.-J. Chale (CHU de Toulouse), P. Metral (CHU Dijon), G. Lairy (CH Corbeil, ANAES),
- des directeurs d'établissement : M. Guesnier (AP-HP), G. Parmentier (Directeur de clinique), R. Le Moigne (CHU de Nantes),
- des personnes qualifiées issues de la Direction des Hôpitaux (A.-C. Bensadon, J. Dubois), de la CNAMTS (M. Marty-ENSM), d'une Agence Régionale de l'Hospitalisation (A. Corvez, ARH Languedoc Roussillon), d'une Direction Régionale des Affaires Sanitaires et Sociales (C. Lucas, DRASS PACA), de la DREES (D. Baubeau, P. Cuneo).

Ce groupe achève la rédaction d'un premier « Guide du bon usage des données médico-administratives pour la comparaison des performances hospitalières » qui sera disponible au début de l'année 2000.

seuil minimal d'activité sur la qualité des soins existent toutefois dans certains champs comme la prise en charge des maladies cardio-vasculaires, la réanimation néonatale, le traitement du SIDA, de certains cancers et un nombre limité d'interventions (cataracte, cholécystectomie, amputation du membre inférieur, opérations intestinales, chirurgie gastrique, anévrisme de l'aorte abdominale).

Les études probantes mettent en avant tantôt le facteur individuel, à savoir l'expérience du chirurgien lui-même, tantôt les paramètres d'organisation : c'est alors le fait d'opérer dans un établissement donné qui est déterminant. Les seuils minimaux sont très variables pour un même acte selon les organisations, et ils diffèrent d'ailleurs selon les pays.

À l'opposé, un certain nombre d'études concordantes montrent également l'existence dans certains domaines de niveaux d'activité à partir desquels la qualité des soins peut décroître : dans certaines disciplines comme la chirurgie cardiaque, la réanimation et la néonatalogie, les problèmes de ressources disponibles et d'organisation collective deviennent plus délicats à gérer lorsque l'activité

5. L'analogie avec les tests de dépistage souligne le fait que ces indicateurs détectent parfois à tort (« faux positifs »), ou, au contraire, ne détectent pas certaines situations anormales (« faux négatifs »), ce qui impose des contrôles par d'autres « tests » (en l'occurrence des investigations complémentaires sur dossiers ou autres), comme c'est la règle en matière de dépistage du SIDA ou du cancer du sein.

6. Pour mettre en évidence une différence de mortalité statistiquement significative entre différents hôpitaux au décours d'interventions pour prothèse totale de hanche, avec un taux de mortalité de 1 % dans un hôpital et de 0,5 % dans l'autre (soit une différence de 50 %), il faut inclure 4 673 patients par hôpital, ce qui nécessite 20 ans de recueil de données pour un hôpital réalisant 230 prothèses par an (P. Ravaud, B. Giraudeau, P.-M. Roux, P. Durieux, « Les indicateurs de mortalité sont-ils de bons indicateurs de qualité des soins », Presse médicale, 1999, n° 28-29 pp. 1604-1609).

E·2

La notion de performance pour les établissements de santé ?

Les publications qui ont mis en avant un « classement » des établissements hospitaliers ne considèrent en fait qu'une seule dimension de la performance, celle de la **qualité clinique ou plutôt de l'efficacité technique** des soins.

● En effet, tous les paramètres utilisés se caractérisent in fine par leur orientation vers la mesure de critères « quasi-industriels », qu'il s'agisse de volume d'activité (nombre d'actes par an), de taux d'erreurs ou de défaillance (décès), de temps de fabrication (durée de séjour, chirurgie ambulatoire), la part de production à l'exportation (notoriété) venant compléter le tableau.

● Si la pertinence de chacun de ces indicateurs apparaît problématique en regard de leur capacité à mesurer la qualité des soins délivrés, leur ensemble aboutit de facto à une représentation de la « performance » excessivement réductrice, au regard des missions de l'hôpital et en particulier de ses missions de service public. Deux études, l'une canadienne portant sur les établissements de santé, l'autre française concernant la poste présentent des similitudes remarquables intéressant les différentes dimensions de la performance de ces deux types d'industries de service. La première dimension de la performance porte sur ses missions dont découlent ses objectifs : ceux d'une institution de soin étant de concourir à l'amélioration de la santé de l'ensemble de la population et pas seulement de sa clientèle. La seconde dimension est celle de la productivité ; la troisième est relative à sa capacité d'adaptation à l'environnement et d'innovation, (qui inclut sa capacité à répondre aux attentes de sa clientèle et à la satisfaire). La quatrième est celle des valeurs communes qui, parce qu'elles produisent du sens, au sein de l'organisation, garantit la bonne marche des structures grâce à l'implication des équipes dans l'accomplissement de leurs missions liées à leur nature de service public. Il se trouve que les données nécessaires tout comme les indicateurs de mesure des dimensions autres que productives, sont rares, en dehors des enquêtes de satisfaction des patients et de l'accréditation qui devrait en partie permettre de mesurer les deux dernières dimensions de la performance.

Ceci implique une vision élargie de la performance conduisant à associer systématiquement aux mesures de productivité, celles concernant l'évaluation de la qualité des services rendus et de leurs coûts. Dans certaines situations, le maintien de l'accessibilité aux soins s'accompagne d'une perte de productivité.

dépasse un certain niveau et/ou les risques d'infections nosocomiales par contamination croisée augmentent.

■ La mortalité hospitalière

Elle fournit également un excellent exemple des nombreuses précautions à prendre avant de pouvoir porter un jugement s'appuyant sur des indicateurs de résultats. En effet, ceux-ci devraient d'abord être considérés comme des outils permettant de suspecter une zone de « non qualité » et donc, attirer l'attention (un peu comme un test de « dépistage »⁵) et non pas comme des éléments permettant de l'affirmer (comme un test diagnostique). Ces indicateurs permettent rarement d'aider à la recherche des causes de dysfonctionnements et de réponses adaptées pour les corriger (encadré 3).

En premier lieu, un taux de mortalité global n'a pas de validité car il traduit un taux « moyen » sans réalité concrète puisqu'il porte sur une

clientèle hétérogène et peut refléter de « bons » et de « mauvais » résultats dans des disciplines différentes. Pour avoir du sens, les études doivent se centrer sur un acte, une pathologie ou une procédure médicale définie, à condition que ceux-ci soient suffisamment fréquents dans les établissements étudiés (une centaine d'observations annuelles au minimum).

Pour mettre en évidence une différence statistiquement significative de mortalité (événement heureusement rare), il est nécessaire de réunir un très grand nombre d'observations⁶, ce qui impose généralement de rassembler les données sur plusieurs années.

Il faut, bien évidemment, « ajuster », c'est-à-dire tenir compte de l'âge, de la pathologie (on peut être opéré de l'intestin pour des diverticules ou pour un cancer, le risque de décès n'est pas le même dans les deux cas), des co-morbidités associées (un malade opéré de la hanche peut être également diabétique ou souffrir



d'une insuffisance cardiaque), ainsi que du contexte de l'acte (l'admission en urgence est statistiquement reliée à une plus grande mortalité) et, dans certains cas, du niveau de gravité des lésions (par exemple le degré d'insuffisance cardiaque). Tous les taux publiés à l'étranger sont, ainsi, ajustés sur de nombreuses variables, différentes suivant l'acte étudié. Les données du programme médicalisé des systèmes d'information (PMSI) ne contenant pas ce type d'information, le retour aux dossiers médicaux est donc indispensable à la conduite de telles études, à la fois pour compléter

les informations et vérifier l'exactitude des éléments contenus dans le résumé standardisé de sortie (RSS).

La mortalité hospitalière peut par ailleurs dépendre de ce qui s'est passé avant l'hospitalisation : la mortalité hospitalière liée à l'asthme, par exemple, est fortement conditionnée par l'accessibilité des services ambulatoires, tout comme les taux de mortalité néonatale et de mortalité maternelle sont liés à la qualité du suivi de la grossesse (un tiers des 67 dossiers de morts maternelles, survenues en 1996 et expertisés, aurait pu être évité par un meilleur suivi de la grossesse⁷).

L'organisation et les pratiques des établissements interviennent aussi, comme dans la durée moyenne de séjour. Certains établissements transfèrent les malades graves dans un autre établissement ou font des « sorties in extremis », alors que d'autres les prennent en charge jusqu'au dernier instant, ce qui fausse bien évidemment les comparaisons. Pour interpréter correctement des taux, il importe de s'intéresser aux modes de sortie des établissements et à ce qui a suivi.

Un dernier élément de confusion est lié à l'absence de robustesse de ces modèles : les classements varient de façon considérable d'une année sur l'autre sans que cela puisse s'expliquer seulement par des modifications de la qualité ; d'autre part, ils diffèrent également selon les sources utilisées. En particulier, les résultats obtenus à partir des bases de données informatisées ne sont pas les mêmes que ceux reposant sur les données plus complètes du dossier.

Pour toutes ces raisons, les auteurs anglo-saxons estiment que les tableaux de bord comparant des taux bruts de mortalité sont trompeurs et qu'il n'est, en général, pas possible de montrer un lien direct entre la qualité technique ou organisationnelle des soins et la valeur de ce taux. Aussi doit-on préférer les mesures de mortalité spécifique (selon les causes) à date fixe : 30, 90 ou 180 jours après l'admission, même si elles sont beaucoup plus difficiles à réaliser, car elles imposent d'organiser un suivi après la sortie et d'étendre les systèmes d'information hors de l'hôpital (assurance maladie, certificats de décès...).

Il faut enfin ajouter que ces taux de mortalité sont inadaptés à l'évaluation de nombreuses activités hospitalières. C'est par exemple le cas pour la psychiatrie, la chirurgie orthopédique réglée, la chirurgie urologique, les soins de suite et de réadaptation, ou les activités médico-techniques.

7. Comité national d'experts sur la mortalité maternelle – Premiers résultats, G. Lévy, M.-H. Bouvier-Colle.

E-3

Les trois types d'indicateurs de qualité

Les indicateurs « de résultats » de santé (taux de mortalité, taux de complications) ont des défauts communs : ils mesurent des événements rares et, dès lors, les tests statistiques manquent de puissance pour pouvoir conclure à une différence entre deux établissements ou à une évolution dans le temps pour un établissement. Ils sont très sensibles à des caractéristiques dites « extrinsèques », c'est-à-dire extérieures à l'établissement et liées au type de recrutement, au statut de l'établissement (universitaire ou non), aux modes de prises en charge (DMS, réadmission, transferts), toutes choses qui imposent un contrôle minutieux des données par retour au dossier ainsi qu'une connaissance fine de l'organisation des établissements. Par ailleurs, ils n'apportent aucun éclairage sur les éventuels dysfonctionnements en cause, et sont largement insuffisants à eux seuls pour entreprendre des actions correctrices. Enfin, du fait de leurs difficultés de maniement dans une optique de comparaison, ils sont mal acceptés par les professionnels de qui dépend, pourtant, la qualité des informations recueillies.

Les indicateurs « de procédures » qui renseignent sur les pratiques et les manières de faire, posent en général moins de problèmes. Tout d'abord, le nombre d'observations nécessaires pour détecter des différences de qualité des soins est considérablement réduit¹ du fait de la meilleure sensibilité des indicateurs (chaque acte est significatif alors que pour l'indicateur de mortalité, seuls les décès le sont), ce qui permet de raccourcir le délai de réaction. Parce qu'ils mesurent directement ce qui a été fait (ou ne l'a pas été), ils donnent des informations utiles pour l'action : si on constate que seulement la moitié des patients justiciables d'un traitement par l'aspirine après un infarctus du myocarde en a reçu, il est possible d'identifier ce qui reste à améliorer. Ils permettent de s'intéresser aux pathologies ne faisant pas appel à des actes techniques, aux activités de prévention et d'éducation sanitaire ainsi qu'aux activités « supports » qui conditionnent la qualité. C'est ainsi qu'on peut interroger les établissements sur la collecte des informations médicales (résultats des contrôles qualité du PMSI), la mesure de la satisfaction des patients (nombre de patients interrogés par questionnaire ou interviews, nombre de plaintes traitées), le bon usage des antibiotiques ou les efforts de formation du personnel destinés à accompagner les politiques nationales en faveur de la lutte contre la douleur ou sur d'autres thèmes relatifs à la sécurité (hygiène, sécurité transfusionnelle et anesthésique, médicament). Enfin, ils sont facilement compréhensibles par les cliniciens.

Les indicateurs « de structures » qui s'intéressent aux ressources humaines, matérielles et financières consacrées aux soins sont quant à eux généralement faciles à obtenir, un certain nombre de données étant disponibles dans les enquêtes administratives comme la Statistique annuelle des établissements (SAE). Bien entendu, le niveau des moyens n'est pas, en lui-même, significatif de la qualité des soins ; cependant la présence de certaines installations et équipements est un facteur qui conditionne la capacité d'un hôpital à remplir une mission donnée. De même, le niveau d'encadrement peut constituer un élément déterminant pour assurer correctement une prestation.

1. Pour détecter une différence de qualité dans la prise en charge de l'infarctus du myocarde, Mant et Hicks ont calculé qu'il fallait 8 179 patients par hôpital si l'on utilise les taux de mortalité et 75 si on utilise des indicateurs de procédures (pourcentage de patients recevant un traitement adapté)- article précité de P. Ravaut et col., Presse Médicale 1999.

■ La « notoriété » ou l'attractivité

Elle est le plus souvent mesurée par le pourcentage des patients traités habitant dans un département autre que celui de l'hôpital. Cet indicateur est discutable pour au moins quatre raisons :

– En premier lieu, ces valeurs d'attractivité sont largement dépendantes de la situation géographique et de l'environnement de l'hôpital. Un établissement situé en bordure de la frontière administrative d'un (ou de deux) départements et bien desservi par la route, apparaît, par exemple, plus « attractif » qu'un autre situé en zone côtière et éloigné des départements limitrophes, sans que ce phénomène n'ait de signification en termes de qualité des soins. De même, plus un établissement est le seul de sa spécialité dans une région, plus il « attire » de patients.

– Certains établissements drainent une population extérieure à leur département et n'attirent que très peu la population du département, en raison d'un phénomène de « sélection » de la clientèle. Un indicateur complémentaire relatif à l'attraction sur la population potentielle du département d'implantation apparaît indispensable à l'analyse.

– L'information sur l'adresse du patient est de qualité encore déficiente, à l'heure actuelle, avec, comme conséquence, de nombreuses valeurs aberrantes liées au fait qu'un certain nombre d'établissements codent « hors département » les informations manquantes. Cette donnée n'étant pas utile au calcul des budgets ne fait en effet pas l'objet de contrôle de qualité. On peut toutefois espérer que cette limite sera transitoire.

– Enfin, les schémas régionaux d'organisation sanitaire tendent désormais à privilégier, pour chaque établissement, une mission différente au sein de « réseaux gradués de prise en charge » aussi bien pour la naissance, les urgences, la cancérologie que pour d'autres thématiques. Tel établissement est surtout appelé à jouer un rôle de proximité, tel autre à intervenir en seconde intention ou à servir de recours. La comparaison des « attractivités » doit dès lors tenir compte de ce positionnement institutionnel différent : pour un hôpital donné, remplir correctement la mission de réponse de proximité qui lui est demandée fait, en effet, partie de sa « performance ». L'utilisation directe des données médico-administratives risque d'attribuer à des différences de « notoriété » des phénomènes liés à l'organisation même des réseaux, ce qui n'a rien à voir avec la qualité technique des soins et pénalise les structures qui jouent le jeu de la coopération et de la complémentarité.

■ La durée moyenne de séjour (DMS)

Cet indicateur peut, dans certains cas, constituer un marqueur de productivité⁸ et de bonne organisation, mais il n'est pas intrinsèquement un indicateur de qualité des soins. Dans les pays où il est souvent utilisé pour les comparaisons, il est complété par le taux de réadmission car une sortie trop précoce, notamment en gériatrie, peut se traduire par une dégradation du service rendu à terme, ce qui a été bien étudié à l'étranger. Il serait d'ailleurs nécessaire de distinguer entre réadmissions programmées et non programmées.

Quoiqu'il en soit, l'analyse correcte de différences de DMS ne peut se faire sans connaître l'organisation des établissements et les habitudes de prise en charge médicale : tel établissement (A) transférera rapidement tous ses patients ayant présenté un pontage coronarien dans un établissement de soins de suite et de réadaptation, tel autre (B) gardera les patients jusqu'à ce qu'ils soient en

mesure de rentrer à leur domicile et de suivre un programme de ré-entraînement à l'effort en ambulatoire. La qualité technique de la prise en charge n'est pas moindre dans le second cas, même si le séjour est plus long. En terme d'efficacité économique globale pour l'assurance maladie, il est même possible que la stratégie de l'hôpital B soit moins coûteuse.

Par ailleurs, l'interprétation de la durée moyenne de séjour doit aussi tenir compte des caractéristiques sociales des patients : les études montrent en effet que les durées de séjour ont tendance à être plus longues lorsque les patients connaissent des conditions de précarité et d'isolement qui rendent difficile leur prise en charge à la sortie de l'hôpital.

Plus le séjour est court, plus le taux de mortalité intra-hospitalière risque d'ailleurs d'être faible et moins il a de sens. Aux États-Unis, une étude⁹ comparant les hôpitaux de la côte Est à ceux de la côte Ouest montrait que les patients hospitalisés dans les hôpitaux de la côte Est avaient un taux de mortalité hospitalière inférieur de 45 % (16,1 %/24,5 %) à ceux de la côte Ouest, avec une durée de séjour trois fois inférieure (8,6 jours/24,5 jours). Or, l'analyse des taux de mortalité à 30 jours (après leur entrée dans l'hôpital) montrait des taux de mortalité identiques ! La différence de « mortalité hospitalière » était due à un artefact statistique lié aux pratiques différentes en matière de sorties et transferts, effet gommé par la prise en compte des morts survenues à l'extérieur de l'hôpital. On aurait donc conclu à tort, si on avait considéré, sur la base de cet indicateur de DMS, que les hôpitaux de l'ouest délivraient des soins de moindre qualité à leurs malades.

■ Les taux d'infections nosocomiales¹⁰

S'ils apparaissent fondamentaux pour les hospitaliers comme pour les usagers, ces taux sont fortement dépendants de plusieurs éléments. Il s'agit d'abord de la qualité du re-

8. La productivité correspond au rapport entre les ressources engagées et la production réalisée.

9. Jenks cité par la Recherche n° 324, octobre 1999, p. 64-78.

10. Selon le texte préparé par le CTIN sur les indicateurs de qualité dans le domaine de l'infection nosocomiale.

cueil : la meilleure façon d'avoir des taux très bas est de ne pas recueillir l'information. Il s'agit ensuite des caractéristiques des patients pris en charge dans l'hôpital ou le service hospitalier. Plus les malades sont fragiles (patients âgés ou nouveaux nés, malades atteints de cancers, immunodéprimés, etc.) et plus ils sont exposés à des actes invasifs¹¹, plus le risque de contracter une infection nosocomiale est élevé. Les critères de définition de certaines infections et la démonstration de leur caractère « acquis » est, souvent, source de difficultés majeures. Une infection est dite nosocomiale si elle apparaît au cours ou à la suite d'une hospitalisation et si elle était absente à l'admission à l'hôpital¹². Ces incertitudes et ces discussions liées au repérage de ces infections rendent impossible d'envisager, pour la plupart d'entre elles, des taux suffisamment fiables pour être comparés comme tels entre services et hôpitaux. Toutefois, les infections du site opératoire peuvent bénéficier d'une méthodologie de recueil et d'ajustement standardisée¹³ et acceptée pour envisager des comparaisons inter-établissements, à condition de vérifier la qualité des données et notamment la durée d'observation. Il faut, en effet, disposer d'un recul d'un mois avant de pouvoir af-

firmer qu'il n'y a pas eu d'infection postopératoire à caractère nosocomial (un an dans le cas de prothèses). Ceci impose un suivi largement au-delà de la date de sortie, ainsi qu'un enregistrement sur des supports complémentaires au système PMSI basé sur le résumé de sortie. La lourdeur de cette procédure et la fragilité des données rendent en outre indispensable un accord et une participation des professionnels (encadré 4) pour que les mesures soient réellement fiables. Ainsi, par exemple, aux États-Unis, la loi protège l'anonymat des établissements qui recueillent ces informations dans le cadre du programme national de lutte contre les infections nosocomiales¹⁴.

■ Les taux de césariennes

La césarienne constituant un geste « à risque » pour la mère, un taux élevé de césariennes est généralement considéré comme un défaut de qualité des soins ; toute la question est donc celle du « bon taux », d'autant qu'on observe d'importantes variations selon les maternités.

Son interprétation nécessite au minimum de tenir compte du niveau de risque obstétrical des femmes prises en charge par les équipes ainsi que celui du fœtus. Le fait que le recueil des données médicales soit effectué

à la sortie sans mentionner les circonstances d'admission, les caractéristiques sociales des mères et les éléments du suivi de la grossesse (encadré 5), font que la plupart des facteurs de risque de césarienne liés aux mères et aux enfants ne sont pas pris en compte dans le PMSI.

Par ailleurs, un décret récent institue une gradation des maternités en trois niveaux : maternité de type 1, 2 ou 3. À chaque niveau correspond une mission, et donc un recrutement particulier, ce qui justifie *a priori* des taux de césarienne différents. Ces structures ne peuvent donc être directement comparées les unes aux autres sans précautions.

Une fois les « groupes homogènes de maternité » constitués (ce qui sera possible en routine à l'échelon national à partir de 2001), il faudra tenir compte du fait qu'à côté des différences de clientèle et de pratique médicale, il existe des différences d'organisation, lesquelles expliquent largement des variations dans le recours aux césariennes : plusieurs études¹⁵ montrent le rôle joué par la disponibilité du personnel d'anesthésie, les attitudes du personnel infirmier, la politique d'entrée en salle de travail ainsi que les attentes des patientes elles-mêmes.

E•4

Une condition de validité : l'appropriation des indicateurs et de la démarche évaluative par les professionnels

Les indicateurs peuvent être définis comme « des données qui permettent de caractériser une situation et d'effectuer des comparaisons dans le temps et l'espace » mais, pour ce faire, ils doivent avoir certaines caractéristiques, en particulier une fiabilité et une validité suffisantes. Lorsqu'ils ne les possèdent pas et, qu'en conséquence, ils ne peuvent être utilisés pour les comparaisons entre établissements, certains indicateurs comme, par exemple, des taux spécifiques d'infections nosocomiales restent utilisables en interne pour suivre l'impact de mesures d'amélioration de la qualité. Ce sont même des outils irremplaçables pour ces démarches au sein des établissements.

En raison de leurs limites, les indicateurs énumérés ont, avant tout, en utilisation externe, un rôle de surveillance et de déclenchement d'investigations plus poussées. Pour jouer pleinement leur rôle d'aide à la décision, il importe que les indicateurs élaborés et proposés pour les comparaisons recueillent l'adhésion des professionnels, et surtout, qu'ils soient utilisés de façon habituelle au sein des établissements pour suivre et rendre compte des démarches d'amélioration de la qualité. Cette appropriation est la condition nécessaire, non seulement de l'amélioration de la réflexion collective, mais aussi de la fiabilité du système d'information qui dépend, pour une large part, de la coopération et de l'implication des professionnels.

11. Ce qui renvoie, là encore, à la question de la pertinence des actes effectués.

12. Définition tirée du document « 100 recommandations pour la surveillance et la prévention des infections nosocomiales » Comité Technique National des Infections Nosocomiales (CTIN).

13. L'indice National Nosocomial Infection Surveillance System (NNIS) est calculé en tenant compte de l'état de santé du patient, du type d'intervention plus ou moins « propre » et de la durée de celle-ci. Il faudra donc comparer des taux d'infection de site opératoire à 30 jours à indice NNIS comparable, en particulier les patients dits NNIS = 0, c'est-à-dire présentant *a priori* un risque faible.

14. Réseau NNIS du Center for Diseases Control d'Atlanta.

15. GAIN Obstétrique CNAMTS-IMAGE.

L'analyse du taux de césariennes au sein de « groupes homogènes de maternités » peut acquérir une signification intéressante à la fois pour les usagers et les décideurs, et devenir, en tant qu'indicateur de procédure, un excellent outil de l'évaluation de la qualité technique et organisationnelle des soins obstétrico-pédiatriques et de leur amélioration, à condition de le compléter par des informations issues des dossiers et du fonctionnement des unités.

« Qualité » et « performance » hospitalières, des notions pluridimensionnelles

La notion de qualité des soins est, on l'a vu, pluridimensionnelle. Elle englobe notamment la notion de la satisfaction des patients et du meilleur coût pour le même résultat. Les attentes des patients, des gestionnaires et de la société en général vis-à-vis des établissements de santé dépassent le soin proprement dit et concernent aussi le service rendu à une population.

– Les établissements doivent s'organiser pour répondre aux besoins de la population, ce qui implique en particulier qu'ils se rendent accessibles à tous ceux qui en ont besoin.

– On attend aussi de l'hôpital qu'il prenne en compte les besoins psychologiques et sociaux de ses patients, qu'il participe à des actions de prévention et d'éducation pour la santé et s'articule harmonieusement avec les autres acteurs de santé de son environnement, notamment au sein de réseaux.

– Les établissements publics ont enfin des missions particulières dans l'accueil des plus démunis et, plus généralement, dans la réduction des inégalités de santé liées aux problèmes sociaux. En l'absence de données sur les caractéristiques sociales des patients dans les résumés standardisés de sortie, leur action dans ce domaine est difficile à mettre en évidence. À défaut, les caractéristiques

socio-économiques des populations des zones d'emploi correspondantes peuvent constituer une variable permettant d'approcher cette dimension.

– L'accueil des urgences, les missions d'enseignement et de recherche figurent enfin parmi les services spécifiquement demandés à certains établissements.

En second lieu, les attentes des usagers sont multiples et probablement variables suivant les points de vue : celles des patients peuvent différer de celles des professionnels ou des gestionnaires. Cette diversité des attentes et des demandes mérite d'être étudiée et si possible prise en compte. Il importe donc d'envisager la performance hospitalière « dans sa globalité » : un hôpital ne saurait par exemple être considéré comme performant s'il ne s'acquitte pas d'une partie de ses missions, ce qui implique d'examiner sur le même plan un panel forcément vaste d'indicateurs (de qualité clinique, de satisfaction, d'accessibilité, etc.). Du fait de leurs missions différentes, les comparai-

sons pour être pertinentes ne peuvent, en outre, porter que sur des établissements dont les caractéristiques et les contraintes sont suffisamment proches...

Des recherches complémentaires doivent être conduites sur les indicateurs susceptibles de rendre compte de ces diverses dimensions de la performance et sur les conditions de leur comparaison, avec, au fur et à mesure, des études de validation à partir des données disponibles.

Les avancées en perspective

– Un premier ensemble de travaux recouvre les démarches entreprises par les professionnels pour mettre au point et tester des indicateurs destinés d'abord à suivre certaines de leurs activités – acte significatif, pathologie bien définie, sous-groupe de patients aux caractéristiques particulières – et, secondairement, à se comparer à des unités similaires. On peut citer les diverses enquêtes conduites sous l'égide des centres interrégio-

E•5

L'hôpital dans son réseau

Une des limites majeures des indicateurs décrits est liée au fait qu'ils isolent l'hôpital de son environnement et ne s'intéressent qu'à un séjour. On voit par exemple que, dans le cas de la naissance, il serait aberrant de s'intéresser uniquement à l'accouchement pour analyser la mortalité périnatale et maternelle. Ce sont les conditions de la grossesse et la qualité du suivi (généralement en consultation) qui sont prépondérantes. Dans d'autres domaines, en raison du raccourcissement des durées de séjours et du développement des alternatives à l'hospitalisation, du fait aussi de la plus grande fréquence des maladies chroniques et du vieillissement de la population, l'épisode de soins que constitue une hospitalisation prend également une place de plus en plus marginale dans le processus de prise en charge. L'image obtenue à partir d'un séjour n'est donc plus à elle seule représentative de la qualité des soins.

De nouveaux indicateurs de qualité apparaissent alors nécessaires pour appréhender les procédures et les résultats obtenus par un ensemble d'acteurs au sein d'une population donnée. Si l'on prend l'exemple du diabète non insulino-dépendant (désormais appelé type 2), un récent travail de l'échelon national du service médical met en évidence des disparités géographiques importantes dans la qualité du suivi, notamment au niveau du dépistage des complications ophtalmologiques (examen spécialisé tous les ans) et cardio-vasculaires (électrocardiogramme de repos). Dans ce contexte, l'incidence des complications ou le taux de couverture des examens de dépistage, au sein de la population de diabétiques desservie par un établissement, sont plus informatives que les données sur les séjours hospitaliers pour diabète. Ils explorent la qualité d'une prise en charge globale (y compris les efforts d'éducation pour la santé) et partagée, ainsi que la capacité d'un établissement à travailler en partenariat pour améliorer le suivi et évaluer les résultats. Ce type d'indicateurs nécessite que les établissements connaissent leur « file active » et peut faire appel à des données extérieures à l'établissement (données de l'assurance maladie sur les affections de longue durée par exemple). Bien entendu, tenir compte des différences de caractéristiques des patients reste déterminant dans ce type de comparaisons.

naux de lutte contre les infections nosocomiales (CCLIN)¹⁶ regroupés en réseau national – Réseau national de surveillance de l'infection nosocomiale (RESIN) ; en matière de prise en charge de la naissance : le réseau Association des utilisateurs de dossiers informatisés en périnatalogie, obstétrique et gynécologie (AUDIPOG), celui du Groupe d'évaluation nationale en néonatalogie et le « tableau de bord » de certains comités techniques régionaux de la naissance ; pour les urgences, la publication du guide des outils d'évaluation par la Société de Réanimation de Langue française (les réanimateurs avaient, auparavant, proposé des indicateurs adaptés à leur activité) ; en chirurgie ambulatoire, les recommandations de l'Association française de chirurgie ambulatoire (AFCA). Ce ne sont là que des exemples parmi beaucoup d'autres. Les fédérations hospitalières développent, de leur côté, des banques de données partagées basées sur le volontariat, qui correspondent au « benchmarking » (littéralement, comparaison avec le meilleur) largement pratiqué dans le monde industriel. Les rapports d'accréditation de l'ANAES qui seront disponibles au public constitueront également des éléments d'information qualitative de nature à informer les usagers, bien que leur objet ne consiste pas à comparer les établissements entre eux.

À l'étranger, ces démarches, dont la vocation est de mobiliser la communauté professionnelle elle-même, ont montré leur efficacité pour aider les établissements à améliorer leurs performances.

– Une démarche de nature statistique peut également tenter d'apporter des éléments de comparaison plus globaux. Dans la mesure où aucun indicateur, même robuste et pertinent, ne peut de façon valide être analysé isolément, elle consisterait à juxtaposer l'ensemble des indicateurs disponibles pour appréhender un aspect donné de la performance (dès lors que

leur validité au moins partielle est reconnue), et à mettre en regard des regroupements d'établissements les plus « homogènes » possibles du point de vue de leur activité, de leur environnement et des contraintes auxquelles ils sont confrontés. Les analyses effectuées seront alors forcément « multidimensionnelles », à la fois pour tenir compte de la variabilité des caractéristiques des établissements mais également de la pluralité des indicateurs utilisables, y compris pour un aspect de performance donné.

Si l'on prend l'exemple du taux d'infection des plaies opératoires chez des patients présentant un niveau de risque équivalent, sa signification est différente selon l'usage qui est fait des antibiotiques. Une des façons d'obtenir un taux faible consiste en effet à administrer préventivement des antibiotiques à tous les patients, y compris dans les situations où ils ne sont pas recommandés, pratique qui a pour inconvénient notable d'accroître la sélection de germes multirésistants¹⁷. Les comparaisons d'établissements ayant une activité chirurgicale équivalente (nature, mission et clientèle) pourraient porter sur ces trois points : respect des recommandations, niveau de résistance aux antibiotiques et taux d'infections du site opératoire, ce qui conduit à les répartir selon plusieurs axes.

Cette démarche statistique nécessite, bien entendu, de rassembler des informations permettant de caractériser les établissements, leur clientèle, leur environnement et leurs performances, à partir d'une analyse de données multicritères. À côté du PMSI, la refonte prévue de la Statistique annuelle des établissements (SAE) devrait y contribuer en améliorant la description des établissements (en particulier, l'organisation du plateau technique et de certaines activités de soins), en complétant les indicateurs disponibles par un certain nombre d'indicateurs de structures et de procédures (comme la part d'anti-

biotiques gérés en dispensation nominative, la part des chimiothérapies préparées par la pharmacie, les moyens mis en œuvre pour prendre en charge la douleur et les soins palliatifs, etc). L'étude des caractéristiques socio-démographiques susceptibles d'influencer le niveau de recours à l'hospitalisation et la morbidité des populations dans les différentes zones d'emploi permettra en outre de tenir compte de leur environnement.

En conclusion, une démarche d'analyse des performances des établissements de santé implique de reconnaître la diversité des attentes qui sont exprimées à leur égard et d'en rendre compte à travers des axes de comparaison multiples. Les indicateurs disponibles doivent en outre être utilisés en les combinant sans pour autant les pondérer, en tenant compte de leurs limites et en comparant des établissements ayant des missions et des environnements suffisamment proches pour que ces comparaisons aient du sens. Le partenariat avec les professionnels apparaît enfin indispensable pour améliorer le système d'information et stimuler les démarches d'amélioration de la qualité mobilisant des indicateurs plus fiables et plus valides. ●

16. Notamment l'enquête nationale de prévalence de mai-juin 1996 et les réseaux d'incidence de certaines infections mis en place par les CCLIN (sites opératoires, infections en réanimation, bactériémies, accidents d'exposition au sang...) portant sur plus de 300 services hospitaliers.

17. Se pose ici la question difficile du bénéfice « individuel » immédiat, versus un éventuel effet délétère plus tardif sur les malades ultérieurs.