

Classification des hôpitaux: nouvel algorithme pour la comparaison entre établissements

Comment comparer entre eux des hôpitaux quant à leur qualité et à leur économicité? Dans ce qui suit, on décrit un algorithme permettant de définir des groupes homogènes sur le plan des coûts et des risques, et donc aptes à servir de base pour une comparaison entre établissements. A la différence de la typologie hospitalière existante de l'Office fédéral de la statistique, le modèle présenté ici est dynamique et peut aussi s'appliquer à des sites hospitaliers individuels.



Kris Haslebacher
Office fédéral de la santé publique

Les prestations et traitements proposés par les hôpitaux doivent à la fois présenter une haute qualité et revenir le moins cher possible. Afin de contrôler dans quelle mesure ces critères sont remplis, les indicateurs

d'économicité et de qualité des différents hôpitaux doivent être comparables entre eux. Pour cela, il faut tenir compte de ce que les hôpitaux ne traitent pas tous la même patientèle, n'offrent pas tous une gamme de prestations semblable et ne disposent pas tous de la même infrastructure. Des comparaisons fondées sur des indicateurs non corrigés statistiquement présentent donc souvent des distorsions systématiques, pouvant conduire à des conclusions erronées.

Pour améliorer la transparence et les possibilités de pilotage du système de santé, il faut donc que les indicateurs des différents hôpitaux soient comparables. L'article qui suit présente une méthode permettant de définir sept catégories d'hôpitaux pour les soins soma-

tiques aigus, à l'intérieur desquelles une comparaison des coûts moyens normalisés par cas¹ est possible².

Contrairement à la typologie hospitalière existante de l'Office fédéral de la statistique (OFS; voir encadré en fin d'article), le modèle nouvellement développé présente l'avantage de s'appliquer aussi bien à des groupes d'hôpitaux qu'à des sites hospitaliers individuels.

Sur la base des trois indicateurs *Casemix-Index (CMI; mesure de la gravité moyenne des cas hospitalisés), nombre de DRG dans l'hôpital (Diagnosis Related Groups, groupes de cas par diagnostic, ci-après: DRG) et proportion de high outliers sur l'ensemble des cas de l'hôpital (proportion de séjours stationnaires dépassant la durée limite, ci-après: PHO)*, il est possible d'assigner chaque hôpital à l'une des sept catégories. Pour cela, les indicateurs de l'hôpital sont comparés à des valeurs seuils calculées préalablement, dans le cadre d'un arbre décisionnel (voir graphique **G1**).

La définition de ces seuils est dynamique, c'est-à-dire que, pour une période donnée, ils peuvent aussi bien être calculés *ex ante* que recalculés *ex post*. Cela garantit, d'une part, que des changements structurels du système hospitalier se reflètent dans la typologie, par le biais des valeurs seuils des indicateurs. D'autre part, cela confère à la méthode une certaine robustesse vis-à-vis d'éventuelles velléités de manipulation de la part des hôpitaux, qui pourraient être tentés d'atteindre tout juste un certain seuil afin de se placer dans une autre catégorie.

L'algorithme présenté ici est le résultat des plusieurs itérations de développement, le but étant de produire une typologie présentant, au sujet de l'économicité (définie opérationnellement comme le coût moyen normalisé par cas), un pouvoir explicatif³ maxi-

1 Coût moyen normalisé par cas = $(1/CMI) \times \text{coût moyen par cas}$

2 Une autre condition pour que la comparaison ait un sens, c'est une certaine homogénéité des coûts à l'intérieur du groupe considéré. Pour la catégorie hospitalière 5 du moins (voir tableau T1 et commentaire), cette condition n'est actuellement pas remplie.

3 Le pouvoir explicatif de cette typologie sur le coût moyen normalisé par cas, avec un R2 ajusté d'environ 0,35 pour l'année 2012, est à peu près égal à celui de la typologie hospitalière de l'OFS. Sources des données pour ces calculs: Statistique médicale des hôpitaux 2012 et Données économiques par cas 2012: www.bfs.admin.ch → Thèmes → 14 - Santé → Prestations et → recours aux services → Hôpitaux → Données détaillées → Tableaux standards → 2010-2013

mal pour un nombre minimal de catégories d'hôpitaux. Cette méthode de classification fondée sur des seuils et un arbre décisionnel mise délibérément sur une logique dont l'application soit aussi simple que possible, une fois les seuils définis.

Outre le choix des indicateurs, l'évaluation de la méthode a porté sur le nombre de seuils nécessaire par indicateur et sur la possibilité de regrouper certaines catégories d'hôpitaux.

L'algorithme

Le but était de développer un algorithme par lequel chaque hôpital n puisse être assigné de façon univoque à une catégorie d'hôpitaux au moyen de M facteurs de catégorisation. Pour cela, on commence par diviser chaque facteur de catégorisation en un certain nombre d'intervalles qui ont du sens et permettront de placer chaque hôpital dans un sous-groupe k pour ce facteur. En combinant les sous-groupes obtenus pour chacun des M facteurs de catégorisation, on obtient une catégorie d'hôpitaux pouvant comporter au maximum $\prod_{m=1}^M K_m$ catégories. En fusionnant certaines des combinaisons possibles, on peut réduire encore ce nombre, pour autant que ces regroupements se justifient par la nature des choses.

La détermination des différents sous-groupes relatifs à un facteur de catégorisation donné s'effectue par la méthode dite *k-means clustering* ou partitionnement en *k-moyennes*⁴. Cette méthode consiste à minimiser la somme des carrés des distances entre les valeurs observées et le centre de gravité de leur cluster (sous-groupe provisoire) respectif. Pour le présent algorithme, ce calcul a été effectué au moyen du logiciel statistique *R* et selon la méthode appliquée par *Hartigan et Wong*⁵.

Pour chaque facteur de catégorisation, on commence à définir un seuil au-delà duquel une observation est considérée comme aberrante

(*outlier*) et n'interviendra pas dans la détermination des sous-groupes. Cette étape est nécessaire, sans quoi la méthode de clustering produit pour certaines variables des sous-groupes ne comportant dans les cas extrêmes qu'un seul hôpital. À l'examen du corpus de données utilisé, il est vite apparu qu'une règle automatique et simple pour identifier les valeurs aberrantes n'aurait rien donné d'utile. Les seuils pour les valeurs aberrantes ont donc été définis heuristiquement.

Calcul des valeurs seuils des différents facteurs de catégorisation

Après convergence de l'algorithme *k-means*, on détermine les limites entre les intervalles contenant les clusters, et ces limites servent ensuite de seuils. La limite entre les intervalles de deux clusters voisins est décrite par la formule suivante :

$$\frac{\left(\begin{array}{l} \text{valeur la plus haute} \\ \text{de celles observées} \\ \text{dans le cluster dont} \\ \text{les valeurs sont} \\ \text{plus basses} \end{array} \right) + \left(\begin{array}{l} \text{valeur la plus} \\ \text{basse de celles} \\ \text{observées dans} \\ \text{le cluster dont} \\ \text{les valeurs sont} \\ \text{plus hautes} \end{array} \right)}{2}$$

- **Seuils pour le CMI:** Pour le CMI, on calcule au total deux valeurs seuils (définissant trois sous-groupes). La procédure de clustering prend en compte les valeurs de tous les hôpitaux, sauf de ceux qui présentaient un CMI > 1,5 (valeurs aberrantes).
- **Seuil pour les DRG:** On calcule ici un seuil unique (ce qui donne deux sous-groupes). La procédure de clustering n'est pas appliquée à l'ensemble des hôpitaux, mais seulement à ceux qui présentent un CMI supérieur au premier seuil de la variable CMI. Ce filtrage correspond à la fusion de certains sous-groupes et correspond à la seconde

étape décisionnelle dans le graphique **G1**.

- **Seuil pour la PHO:** Pour la PHO également, on calcule un unique seuil (d'où deux sous-groupes). Après exclusion des *outliers* (PHO > 0,15), la procédure de clustering ne prend en compte que les valeurs des hôpitaux qui :
 - présentent un CMI supérieur au premier seuil de la variable CMI;
 - traitent davantage de DRG différents que le seuil défini pour la variable DRG.

Ce filtrage obéit lui aussi à la logique de la fusion de sous-groupes.

Fusions de sous-groupes

Une fois calculés tous les seuils, on obtient les sept catégories d'hôpitaux, en procédant donc selon le schéma du graphique **G1**. Pour clarifier la situation, prenons un petit exemple hypothétique. Supposons que les valeurs seuils des trois facteurs de catégorisation ont été calculées *ex ante* et que l'on a obtenu les valeurs suivantes :

- Seuils supposés :
- CMI 0,60 ; 0,97
 - DRG 360
 - PHO 0,05

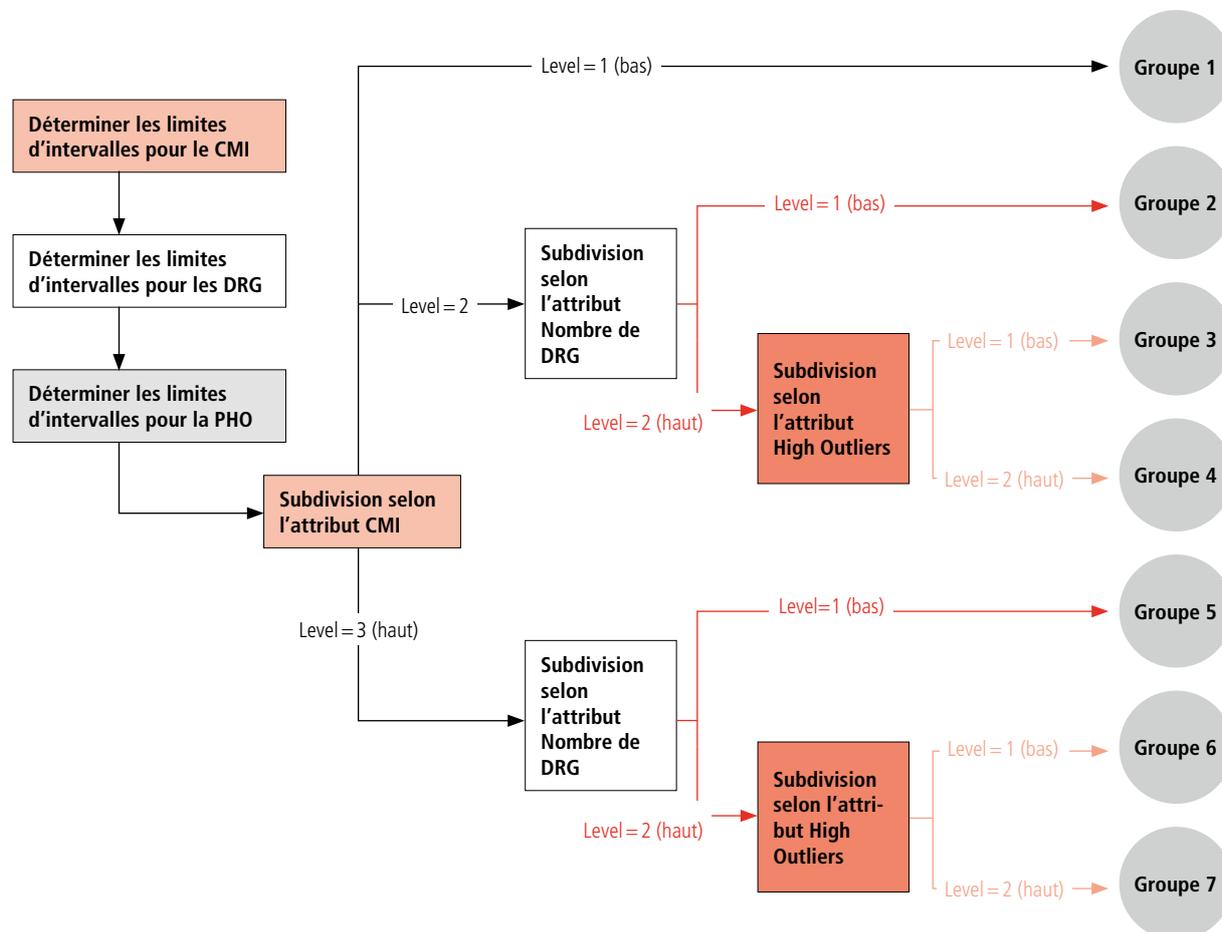
L'hôpital hypothétique n présente un CMI de 0,71, 402 DRG différents et une PHO de 2%. Dans le classement selon l'attribut CMI, cet hôpital se place entre le premier et le second seuil, c'est-à-dire dans le sous-groupe du milieu (Level 2), ce qui signifie que l'on suit la branche moyenne de l'arbre décisionnel. Ensuite, on compare l'attribut DRG de l'hôpital avec la valeur seuil. Comme 402 est supé-

4 Comme centroïdes initiaux pour amorcer l'algorithme, on a choisi les moyennes des sous-groupes définis par les quantiles. Pour $k = 2$, les deux sous-groupes dont on a calculé la moyenne respective étaient donc séparés par la médiane; pour $k = 3$, les sous-groupes étaient séparés par les tertiles, pour $k = 4$ par les quartiles, etc.

5 Hartigan, John A. et Manchek A. Wong, « A k-means clustering algorithm », dans *Applied Statistics* 28, 1979, pp. 100-108.

Arbre décisionnel pour la catégorisation des hôpitaux de soins aigus

G1



rieur à 360, on suit la branche inférieure de l'arbre décisionnel (Level 2). Pour le troisième attribut, on a $2\% < 5\%$, c'est-à-dire que notre hôpital hypothétique se place dans la catégorie 3.

Résultats

A titre de test, sur la base des données de la « Statistique médicale des hôpitaux 2012 » de l'OFS,⁶ on a calculé les seuils à l'aide de l'algorithme présenté ici, en se limitant aux hôpi-

taux de soins aigus. Les seuils obtenus étaient :

- CMI 0,60; 0,97
- DRG 360
- PHO 0,05

Ces seuils ont ensuite pu être utilisés pour ranger dans une des catégories tous les hôpitaux (anonymisés) figurant dans la statistique des « Données économiques par cas 2012 » de l'OFS. A la différence de la Statistique médicale, la statistique par cas ne considère que les données d'environ la moitié des hôpitaux de soins aigus. En revanche, elle contient des indications sur les coûts par cas effectifs, ce qui permet de présenter des chiffres-clés par catégorie d'hôpitaux (voir tableau T1).

Les coûts par cas augmentent de manière continue de la catégorie 1 à la catégorie 7. Fait exception la catégorie 5, dont les hôpitaux présentent un CMI élevé mais seulement très peu de DRG. Par sa définition, cette catégorie comprend de nombreuses cliniques spécialisées offrant des gammes de prestations très diverses, dont l'algorithme présenté ici ne permet pas une prise en compte suffisante. Ce fait se reflète également dans la valeur plus élevée du coefficient de dispersion des quartiles pour cette catégorie. Les coûts par cas des établissements de cette catégorie sont moins uniformes qu'ils ne le sont dans les autres catégories. La valeur la plus élevée des coûts par cas se trouve dans la catégorie 7, qui correspond pour l'essentiel aux hôpitaux universitaires.

⁶ Les facteurs de catégorisation nécessaires à l'algorithme ont été obtenus en groupant les cas à l'aide du SwissDRG-Groupier 2012 V 1.0.

Statistique descriptive par catégorie d'hôpitaux

T1

Catégorie d'hôpitaux	Nombre d'hôpitaux	Médiane des CMNC*	Coeff. de dispersion des quartiles**
1	8	7232	0,18
2	37	8734	0,12
3	31	8735	0,08
4	5	8834	0,15
5	9	10323	0,42
6	5	9382	0,12
7	5	11351	0,12

* Médiane des coûts moyens normalisés par cas, en francs, sur les hôpitaux de la catégorie

** Quantification de l'hétérogénéité des coûts = (3^e quartile – 1^{er} quartile)/médiane

Source: calculs de l'auteur.

Utilisation

L'OFSP utilise la méthode exposée dans le graphique G1 pour ses comparaisons entre établissements. Les bases de données nécessaires à cet effet sont mises à disposition par l'OFS. L'OFS soutient la démarche, il tient à jour et calcule les données de base pour la catégorisation des hôpitaux. L'OFSP et l'OFS se concertent sur la méthodologie et sur son perfectionnement.

Kris Haslebacher, MSc Economics, économiste de la santé, Section Gestion des données et statistique, OFSP.

Mél: kris.haslebacher@bag.admin.ch

La typologie de l'OFS

Contrairement à la typologie présentée ici, celle de l'OFS ne met pas l'accent sur la comparaison entre établissements, mais sur l'ensemble de la gamme des prestations dans le domaine stationnaire. Aussi n'englobe-t-elle pas seulement les hôpitaux de soins somatiques aigus, mais aussi les cliniques psychiatriques et de réadaptation.

La typologie de l'OFS est employée surtout pour permettre de regrouper les établissements de structure comparable dans les statistiques sur les services de santé. Cette utilisation peut se faire aussi bien dans une perspective transversale que longitudinale. En conséquence, la typologie OFS doit être aussi stable que possible sur la durée. D'autres critères importants sont la compatibilité avec des nomenclatures statistiques nationales et internationales ainsi que le caractère simple et compréhensible des différentes catégories.

Depuis l'introduction de la méthode en 2006, la typologie de l'OFS se base sur les caractères suivants: l'étendue de l'offre de prestations, la spécialisation des prestations, la pondération des catégories de perfectionnement professionnel proposées ainsi que le volume total des prestations.

La typologie hospitalière actuelle de l'OFS comprend les catégories suivantes:

K	Hôpitaux	K1	Hôpitaux de soins généraux	K11	Prise en charge centralisée	K111	Niveau de prestations 1 (hôpitaux universitaires)
						K112	Niveau de prestations 2
				K12	Soins de base	K121	Niveau de prestations 3
						K122	Niveau de prestations 4
						K123	Niveau de prestations 5
		K2	Cliniques spécialisées	K21	Cliniques psychiatriques	K211	Niveau de prestations 1
						K212	Niveau de prestations 2
				K22	Cliniques de réadaptation	K221	Cliniques de réadaptation
				K23	Autres cliniques spécialisées	K231	Chirurgie
						K232	Gynécologie/néonatalogie
						K233	Pédiatrie
						K234	Gériatrie
						K235	Cliniques spécialisées diverses