



**Agence Technique de l'Information sur l'Hospitalisation**

SIRET n° 180 092 298 00033 – code APE : 8411 Z

117, boulevard Marius Vivier Merle

69329 LYON CEDEX 03

Téléphone : 04 37 91 33 10

Télécopie : 04 37 91 33 67

Service Valorisation et synthèse

Pôle Campagnes budgétaires et tarifaires

# **Modalités de calcul du référentiel national de coûts 2007**

## **Données ENCC 2007**

Ce document présente les modalités de calcul du référentiel national de coûts constitué à partir des données ENCC relatives à l'activité 2007 des établissements.

*Avertissement* : les coûts complets ne sont pas directement comparables aux tarifs, compte tenu des différences de périmètre existant entre ces deux concepts. Les données de coûts des référentiels sont destinées essentiellement à permettre le positionnement de l'activité MCO de tout établissement par rapport à l'activité MCO d'un établissement « moyen ».

## **1 – Corrections préalables aux calculs de coûts moyens des séjours**

Les modalités de calcul du coût des séjours prennent en compte certaines corrections préalables, exactement comme les années précédentes :

- Les coûts des séjours des établissements bénéficiant de crédits relatifs au financement de l'enseignement, de la recherche et du rôle de recours sont minorés sur la base de taux d'abattement spécifiques correspondant aux taux MERRI fixés par le ministère. Ils se situent entre 9% et 17%. Ainsi, chaque séjour de la base ENC voit, le cas échéant, son coût minoré du taux de financement des MERRI dont l'établissement a bénéficié. Cet abattement concerne tous les postes de dépenses hors éléments facturés en sus (DMI et médicaments facturés en sus, médicaments ATU, dépenses de SMUR, dépenses de laboratoires Hors Nomenclature, dialyse et radiothérapie en sus, honoraires des PH).
- Dans la mesure où certaines zones géographiques donnent lieu à des surcoûts dont on ne souhaite pas répercuter l'effet dans le référentiel national, les coûts des séjours des établissements concernés ont été diminués à l'aide du coefficient géographique correspondant (la composition de l'échantillon 2007 conduit à minorer ainsi de 7% les coûts des seuls établissements de la région Ile-de-France).

## **2 – Calculs de coûts moyens par GHS ou GHM et opérations de calage**

### A- Opérations de calage

L'ENC est une enquête par sondage auprès d'un échantillon d'établissements, et donc de séjours. De ce fait, les coûts moyens par GHS ou par GHM ne sont que des estimateurs de « vraies valeurs » inconnues. Les coûts moyens calculés sont sensibles à la composition de l'échantillon d'établissements. Cette variabilité, qui se mesure, peut être réduite en utilisant une technique de redressement (dit aussi calage).

Pour résumer, le principe du redressement consiste à tirer profit d'une information auxiliaire – c'est-à-dire en pratique d'un ensemble de variables corrélées au coût complet - disponible dans un fichier national afin de modifier les poids de sondage initiaux. Les poids calés ont pour propriété, d'une part d'être proches des poids initiaux (lesquels reflètent directement la méthode d'échantillonnage) et d'autre part de produire des estimations de totaux, pour chacune des variables auxiliaires, égales aux vrais totaux de ces variables, cela quel que soit l'échantillon tiré. Les vrais totaux sont obtenus à partir d'un fichier exhaustif, en l'occurrence la base nationale annuelle regroupant tous les séjours du PMSI. Le calage permet donc une certaine stabilisation des estimations de coût moyen, cela d'autant plus que les variables auxiliaires sont corrélées au coût complet.

L'individu statistique traité est toujours le séjour : de ce fait, chaque séjour de la base a un poids calé qui lui est propre. Pour réaliser les calages simultanément sur plusieurs variables auxiliaires, le logiciel CALMAR (CALage sur MARGes) développé par l'Insee a été utilisé. Les pondérations calées produites par le logiciel sont telles que :

- pour une variable auxiliaire qualitative (comme le type d'établissement), les effectifs pondérés des modalités de la variable dans l'échantillon, après redressement, seront égaux aux effectifs connus sur la population ;
- pour une variable auxiliaire quantitative (comme la durée de séjour), le total pondéré de la variable dans l'échantillon, après redressement, sera égal au total connu sur la population.

Les bases nationales annuelles de PMSI 2007 ont été utilisées pour calculer les vrais totaux sur lesquels on se cale.

**Pour les établissements ex-DGF**, les variables de calage sélectionnées sont les suivantes :

- **type de l'établissement (5 modalités)**
  - o centre hospitalier universitaire
  - o centre hospitalier de moins de 16 000 RSA annuels
  - o centre hospitalier de plus de 16 000 RSA annuels
  - o centre de lutte contre le cancer (CLCC)
  - o établissement privé participant au service public hospitalier, hors CLCC
- **durée totale du séjour par type d'établissement** (soit 5 variables en réalité, une par type d'établissement)
- **durée du passage en réanimation (= nombre de suppléments REA)**
- **nombre d'actes réalisés (= nombre de zones d'actes dans le RSA)**

**Pour les établissements ex-OQN**, les variables de calage sélectionnées sont les suivantes :

- durée totale du séjour
- durée du passage en réanimation, soins intensifs et soins continus (somme des suppléments REA, STF, SRC et SRA issus du RSF)
- nombre d'actes réalisés (=nombre de zones d'actes dans le RSA).

Le calage s'effectue GHM par GHM (ou GHS par GHS), pour les GHM présentant un nombre suffisant de séjours (au minimum 30 séjours). Cela étant, des aménagements selon le type de GHM sont opérés : ainsi, pour les GHM de courte durée (J et T), la durée du passage en réanimation n'est pas retenue. Pour les GHM de séances, seul le type d'établissement est utilisé pour caler.

La procédure de calage permet de caler sur le plus grand nombre de variables, en sélectionnant les variables pertinentes pour le GHM. Par exemple, si un GHM n'a pas de réanimation, la procédure de calage supprimera cette variable, et le GHM sera donc calé uniquement sur le type d'établissement (pour un établissement ex-DGF), la durée de séjour et le nombre d'actes.

In fine, le coût moyen estimé par GHM (ou GHS) disponible dans le référentiel est issu du calcul suivant :

$$\frac{\sum_{\text{séjours}} \text{Poids redressé du séjour} \times \text{coût du séjour}}{\sum_{\text{séjours}} \text{poids redressé}}$$

## B- Procédures statistiques

- *Trimages en DMS et en coût*

Avant de calculer un coût moyen par GHM, un trimage en DMS et un trimage en coût sont appliqués au préalable.

Le trimage en DMS est appliqué sur les seuls GHM de niveau 1. Les séjours de niveau 1 pour lesquels la durée de séjour est supérieure à 2,5 fois la DMS nationale, hors séjours de 0 ou 1 jour, sont supprimés. En effet, il est supposé que dans ces niveaux, sont présents des séjours à tort, par défaut de codage des DAS (Diagnostics Associés). Dès lors que la v11 sera effective, ces diagnostics devraient être renseignés avec pour effet un classement des séjours dans les niveaux supérieurs. Ainsi, 0,7% des séjours de la base ex-DGF sont trimés et 0,4% dans la base ex-OQN.

Le trimage en coût, consiste à éliminer tous les séjours « extrêmes » dont les coûts sont en deçà d'une borne inférieure ou au-delà d'une borne supérieure définies de la façon suivante :

Soit :  $LC$  : logarithme du coût complet

$$\begin{aligned} borne\_sup &= \text{Exp}(\text{moyenne}(LC) + \text{écartype}(LC) + \text{centil\_95\%}(LC) - \text{médiane}(LC)) \\ borne\_inf &= \text{Exp}(\text{moyenne}(LC) - \text{écartype}(LC) + \text{centil\_5\%}(LC) - \text{médiane}(LC)) \end{aligned}$$

avec,  $borne\_sup$  la borne supérieure de trimage et  $borne\_inf$  la borne inférieure de trimage.

Ce trimage ne s'applique qu'aux GHM de moins de 80 séjours. En effet, les GHM ayant un effectif important de séjours sont moins sensibles aux coûts extrêmes. De plus, les séjours ne sont trimés que lorsque l'écart entre le coût moyen calculé sur tous les séjours et le coût moyen calculé hors séjours extrêmes est supérieur à 10%. Pour le secteur ex-DGF, 67 séjours sont éliminés et 25 séjours dans le secteur ex-OQN.

- *Procédure d'estimation pour les GHM à effectif insuffisant*

Pour les GHM de moins de 30 séjours, le coût n'étant pas calé par CALMAR, une autre procédure d'estimation statistique est utilisée. Cette procédure utilise la hiérarchie des racines de la version 11 des GHM et permet ainsi d'estimer un niveau à partir du niveau inférieur ou supérieur. Par exemple, si un GHM de niveau 3 présente moins de 30 séjours, le coût affiché sera estimé à partir du coût niveau 2 et d'un ratio de coût calculé par sous-CMD. Les coûts ainsi affichés dans le référentiel respectent la hiérarchie des niveaux.

### **3 – Coût moyen des unités d'œuvre**

Pour chaque établissement de l'échantillon et pour une unité d'œuvre donnée, est calculé le coût total dans l'établissement associé à cette unité d'œuvre ainsi que le nombre total d'unités d'œuvre consommées dans l'établissement. Le coût moyen national estimé de l'unité d'œuvre est alors défini comme le ratio (pour le secteur ex-DGF) :

$$\frac{\sum_{\text{établt}} \text{poids} \times \text{cout total établissement}}{\sum_{\text{établt}} \text{poids} \times \text{nombre total UO établissement}}$$

Les sommes concernent l'ensemble des établissements de l'échantillon. Les poids sont des poids de sondage d'établissements, définis comme les inverses des taux de sondage des établissements par type d'établissement.

La formule diffère dans le secteur ex-OQN, du fait de l'absence de type d'établissement :

$$\frac{\sum_{\text{établi}} \text{cout total établissement}}{\sum_{\text{établi}} \text{nombre total UO établissement}}$$

Le référentiel fournit également le nombre total de sections d'analyse (SA) mobilisées pour le calcul ainsi qu'une appréciation de la dispersion des coûts de l'unité d'œuvre. Cette dispersion est définie comme l'écart type des valeurs de coût moyen de l'unité d'œuvre obtenues par section d'analyse (c'est-à-dire le ratio du coût total de l'unité d'œuvre dans la SA par le nombre total d'unités d'œuvre consommées dans la SA).

## 4 – Consommation moyenne des unités d'œuvre

Pour chaque GHM, est calculée la consommation moyenne d'unités d'œuvre en utilisant les poids de sondage des séjours concernés. L'estimation prend la forme :

$$\frac{\sum_{\substack{\text{Séjours} \\ \text{tirés}}} \text{Poids redressé du séjour} \times \text{nombre d'unités d'oeuvre du séjour}}{\sum_{\text{séjours tirés}} \text{poids redressé du séjour}}$$

## 5 – Calculs de précision

L'ENCC produit des estimations de coûts moyens par GHM à partir d'un échantillon d'établissements – et donc d'un échantillon de séjours. L'estimation obtenue est par nature sensible à la composition de l'échantillon. Il est possible, en considérant que les établissements sont issus d'un tirage aléatoire (en réalité, le processus de sélection effectivement appliqué est assimilable à un échantillonnage de volontaires), de mesurer l'ampleur de cette sensibilité. Pour ce faire, la colonne intitulée « Erreur relative d'échantillonnage » de l'onglet « intervalle de confiance » fournit l'erreur relative (ERE) de l'estimateur du coût moyen par GHM due à l'aléa d'échantillonnage des établissements, après avoir éliminé les séjours extrêmes. Ce calcul d'erreur prend en compte la procédure de calage exposée dans la partie 3, dont l'objectif spécifique est de limiter l'incertitude liée à l'échantillonnage.

Le calcul d'erreur va prendre en compte les variables utilisées dans le calage des poids au travers des résidus  $U_{ij}$  (avec  $i$  l'établissement et  $j$  le nombre de séjours de l'établissement) de la régression multiple des coûts complets  $C_{ij}$  sur l'ensemble des variables de calage (où  $i$  correspond à l'établissement et  $j$  correspond au séjour). La régression est conçue GHM par GHM. Pour un GHM donné, la régression est effectuée sur l'ensemble des séjours de l'échantillon de l'ENC (tous les types apparaissent donc dans la même régression lorsqu'il s'agit des ex-DGF).

$$\text{On pose } U_i = \sum_{j \in i} U_{ij} \text{ et } \bar{U}_h = \frac{1}{m_h} \sum_{i \in h} U_i .$$

où :  $m_h$  est le nombre d'établissements du type  $h$  dans l'échantillon de l'ENC 2007,

La variance estimée du coût moyen se formule de la façon suivante :

$$\hat{V} = \frac{1}{N^2} \sum_h \left( 1 - \frac{m_h}{M_h} \right) \frac{M_h^2}{m_h(m_h - 1)} \sum_{i \in h} (U_i - \bar{U}_h)^2$$

où :  
 $N$  est le nombre de séjours par GHM dans la base nationale de 2007,  
 $m_h$  est le nombre d'établissements du type  $h$  dans l'échantillon de l'ENC 2007,  
 $M_h$  est le nombre d'établissements du type  $h$  dans la base nationale 2007.

Si un établissement donné ne produit pas de séjour classé dans le GHM considéré, alors il suffit de considérer que  $U_i = 0$ , mais cet établissement participe bien au calcul de la variance.

Il est important de noter que les variables de régression mobilisées pour le calcul des résidus  $U_{ij}$  ne sont pas les mêmes pour tous les GHM puisque les variables de calage sont propres à chaque GHM.

Pour le secteur ex-OQN, la formule diffère du fait de l'absence de type d'établissement. La formule est donc la suivante :

$$\hat{V} = \frac{1}{N^2} \left( 1 - \frac{m}{M} \right) \frac{M^2}{m(m - 1)} \sum_i (U_i - \bar{U})^2$$

où :  
 $N$  est le nombre de séjours par GHM dans la base nationale de 2007,  
 $m$  est le nombre d'établissements dans l'échantillon de l'ENC 2007,  
 $M$  est le nombre d'établissements dans la base nationale 2007.

Pour chaque GHM l'ERE est définie comme le rapport entre l'écart type associé à l'erreur d'échantillonnage (noté  $\hat{V}$  au dessus) et le coût moyen  $\hat{C}$  estimé à partir des données 2007. Exprimée en pourcentage, l'ERE permet d'apprécier immédiatement l'incertitude relative qui entoure l'estimation du coût moyen du GHM.

En ordre de grandeur, il peut être considéré qu'il y a environ 95 chances sur 100 pour que le vrai coût moyen – celui qui serait par définition obtenu à partir de l'ensemble des séjours produits par l'ensemble des établissements du champ MCO – soit compris dans l'intervalle de confiance

$$[\hat{C}(1 - 2.ERE) ; \hat{C}(1 + 2.ERE)]$$

*Exemple :*

Le coût moyen estimé pour un GHM est 5 200 euros et l'ERE calculée vaut 5%. Il y a environ 95 chances sur 100 pour que le 'vrai' coût moyen soit compris entre 5 200.(1 - 2\*5%) euros et 5 200.(1+2\*5%) euros, soit un intervalle de confiance estimé (en euros) égal à [4 680 ; 5 720].

Pour les GHM de moins de 30 séjours ou lorsque les séjours d'un GHM sont produits par un seul établissement, on ne peut pas estimer d'erreur d'échantillonnage : dans ce cas, on mentionne *NC* (non calculé) dans la colonne concernée.