

Guide technique des modalités de calcul du référentiel national de coûts SSR

Données ENC 2016

Service : Financement et analyse économique

Septembre 2018



Table Des Matières

Préambule : La classification en GME	4
1. Opération préalable aux calculs de coûts moyens par GME	5
2. Suppression des observations atypiques	5
2.1 SUPPRESSION DES SEJOURS INCOMPLETS	5
2.2 SUPPRESSION DES SEJOURS/RHA EN ERREUR	5
2.3 SUPPRESSION DES SEJOURS/RHA DITS ATYPIQUES.....	5
3. Calculs de coûts moyens par GME et opérations de calage	6
3.1 OPERATIONS DE CALAGE	6
3.2 CALCULS DE PRECISION	8
3.3 INDICATEUR DE FIABILITE	9

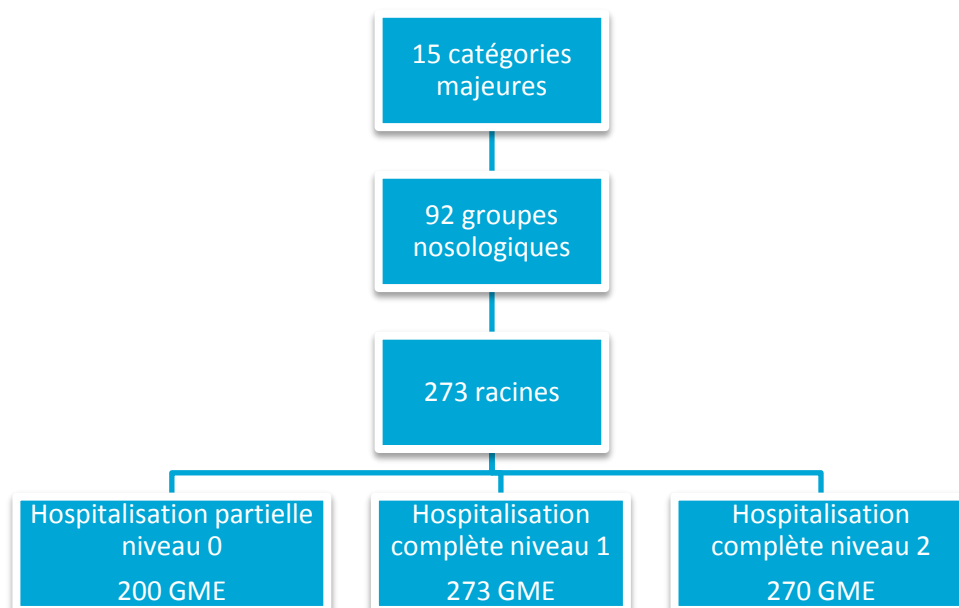
Les données présentées dans le référentiel sont groupées selon la classification en GME (Groupes Médico-Economiques) en vigueur en 2018 (v2018). Le coût moyen par GME est calculé, pour chacun des deux secteurs.

Les principales nouveautés des fichiers Excel du référentiel 2016 :

- Une **nouvelle variable** concerne les charges directes de sous traitance pour les transports (hors SMUR) sur les deux secteurs de financement : *CD Sous-traitance transport (hors SMUR)*.

Préambule : La classification en GME

La version 2018 de la classification est fondée sur 743 GME et se décline en :



L'unité d'observation est soit :

- Le séjour (suite de RHA séparés de moins de 4 jours) en hospitalisation complète¹.
- Le RHA en hospitalisation partielle.

Le référentiel de coûts SSR donne une estimation du coût moyen par GME, soit pour chaque racine selon les catégories suivantes :

- Niveau de sévérité 0
- Niveau de sévérité 1
- Niveau de sévérité 2

Remarque : La distinction Hospitalisation complète / Hospitalisation à temps partiel est réalisée à partir de la variable *Type d'hospitalisation du PMSI*. Néanmoins il n'y a pas toujours adéquation entre les niveaux de sévérités et le type de prise en charge (HP et HC) :

- Le niveau 0 correspond presque exclusivement à de l'hospitalisation partielle, sauf pour des séjours d'hospitalisation complète de 1 jour qui sont groupés dans le niveau 0 quand il existe.
- Les niveaux 1 et 2 correspondent presque exclusivement à de l'hospitalisation complète, sauf pour des journées d'hospitalisation partielle qui sont groupés dans le niveau 1 quand le niveau 0 n'existe pas.

Pour simplifier la lecture dans les différents documents, l'hospitalisation partielle correspond au niveau 0 et l'hospitalisation complète aux niveaux 1 et 2.

¹ Le séjour est donc ici différent du séjour administratif.

1. Opération préalable aux calculs de coûts moyens par GME

Dans la mesure où certaines zones géographiques donnent lieu à des surcoûts dont il convient de ne pas répercuter l'effet, les coûts des séjours/RHA des établissements concernés sont diminués du coefficient géographique correspondant. La composition de l'échantillon 2016 conduit ainsi à minorer de 7% les coûts des établissements de la région Ile-de-France, de 26% les coûts des établissements de la Martinique et de 31% les coûts des établissements de la Réunion.

2. Suppression des observations atypiques

Préalablement au calcul des coûts, une analyse fine des données est réalisée pouvant conduire à la suppression de certaines observations. En effet, il s'agit de ne pas impacter les valeurs moyennes par la prise en compte d'erreurs manifestes.

2.1 Suppression des séjours incomplets

Le référentiel 2016 est construit à partir des informations de l'ensemble des séjours/RHA **complets et terminés en 2016**. Il prend ainsi en compte les séjours à cheval sur 2015-2016 (pour les établissements intégrés aux deux études).

En pratique, pour les séjours commencés en 2015 et terminés en 2016, les coûts des RHA remontés dans l'étude 2015 ont été sommés aux coûts des RHA remontés dans l'étude 2016. Cette opération a été réalisée pour les établissements intégrés aux deux études, soit 55 établissements sur les 62 établissements intégrés à l'échantillon 2016. Les coûts de toutes les sections ont été sommés, à l'exception de la section « accueil et gestion des malades ». Pour celle-ci, seul le coût le plus récent a été conservé (soit le coût de 2016), afin de ne pas imputer deux fois cette charge sur le séjour.

Dans le cas où l'établissement a intégré l'ENC à partir de 2016, ses séjours à cheval sur N-1 ne sont pas pris en compte, faute de connaissance des coûts N-1.

Les séjours à cheval sur trois années (commencés en N-2) ne sont plus pris en compte depuis le référentiel 2015 : ils ne représentaient que très peu de séjours.

Les séjours/RHA à cheval sur N+1 sont eux exclus du référentiel 2016, ils seront pris en compte dans le référentiel 2017 (pour les établissements intégrés en 2016 et 2017).

Les séjours avec des RHA manquants ne sont également pas pris en compte dans le référentiel, faute de connaissance du séjour entier.

2.2 Suppression des séjours/RHA en erreur

Les séjours et RHA groupés en erreur (Catégorie Majeure 90) sont exclus.

2.3 Suppression des séjours/RHA dits atypiques

Sont supprimés quelques séjours/RHA avec les atypies suivantes :

- un coût journalier élevé de transport (RHA d'hospitalisation partielle uniquement)
- une part élevée d'Accueil et gestion des malades (plus de 80% du coût)
- un coût très élevé ou très faible et une incohérence médicale
- un temps métier ou plateau de RR très élevé (plus de 30h/jour)

Au final, les observations supprimées représentent 10% des séjours d'hospitalisation complète (dont 9.5% sont des séjours incomplets) et 1% des RHA d'hospitalisation partielle.
--

3. Calculs de coûts moyens par GME et opérations de calage

3.1 Opérations de calage

L'ENC est une enquête par sondage auprès d'un échantillon d'établissements, et donc de RHA ou de séjours selon le type d'hospitalisation. De ce fait, les coûts moyens par GME ne sont que des estimations du coût national inconnu. Ces coûts sont donc sensibles à la composition de l'échantillon d'établissements. Cette variabilité peut être réduite en utilisant une technique de redressement (appelée aussi calage).

Pour résumer, le principe du redressement consiste à tirer profit d'une information auxiliaire – c'est-à-dire d'un ensemble de variables corrélées au coût – disponible sur l'ensemble des établissements 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 calculés sur l'ensemble de l'activité des établissements à partir de la base nationale PMSI 2016. Le calage permet donc une certaine performance des estimations de coût moyen, et ce d'autant plus si les variables auxiliaires sont corrélées au coût.

L'individu statistique traité est le séjour (suite de RHA) pour l'hospitalisation complète et la journée pour l'hospitalisation partielle. Ainsi, chaque séjour/journée de la base a un poids calé qui lui est propre. Pour réaliser les calages simultanément sur plusieurs variables auxiliaires, la macro SAS CALMAR (CALage sur MARges) développée par l'Insee est utilisée.

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 le score de RR), le total pondéré de la variable dans l'échantillon, après redressement, sera égal au total connu sur la population.

Depuis le référentiel 2013, le calage a été affiné en choisissant les variables de calage au niveau GME (et non plus sur l'ensemble de l'activité). Ainsi, le calage est effectué uniquement sur les variables les plus explicatives du coût pour chaque GME. Par exemple, l'âge n'est plus utilisé automatiquement pour tous les GME, mais uniquement pour ceux dont l'âge est explicatif de la variabilité du coût².

Cet affinement méthodologique permet notamment d'introduire de nouvelles variables « spécifiques », qui sont pertinentes uniquement pour un petit nombre de GME.

Pour chaque type d'hospitalisation, les variables de calage potentiellement utilisées sont les suivantes :

Variables de calage	GME d'hospitalisation complète	GME d'hospitalisation partielle
type de l'établissement	X	X
taux d'occupation de l'établissement	X	
score de dépendance physique	X	X
score de dépendance cognitive	X	X

² Cf. présentation et compte-rendu du comité technique ENC Interchamps du 11 juin 2015, disponibles sur le site internet de l'ATIH

Variables de calage	GME d'hospitalisation complète	GME d'hospitalisation partielle
âge	X	X
nombre d'actes CCAM	X	
score de RR (Rééducation-réadaptation)	X	X
nombre de journées de présence	X	
caractère post-chirurgical de la prise en charge		X
mention spécialisée	X	X
type d'hospitalisation (séance ou HP)		X
mode d'entrée	X	X
mode de sortie	X	

Le type d'établissement :

Sur le secteur DAF, il correspond au statut juridique :

- centre hospitalier universitaire (CHU)
- centre hospitalier (CH)
- établissement privé à but non lucratif (EBNL)

Nouveauté 2016 :

Sur le secteur OQN, il est défini non plus à partir de la taille mais à partir du statut juridique de l'établissement :

- Etablissement privé à but lucratif
- Etablissement privé à but non lucratif

Le taux d'occupation de l'établissement :

Il est issu de la SAE (Statistique Annuelle des Etablissements de santé) 2016, et correspond au nombre de journées réalisées sur le nombre de journées exploitables. Ces données n'étant disponibles que pour l'hospitalisation complète, le taux d'occupation ne peut pas être calculé sur l'hospitalisation partielle.

Trois modalités sont définies pour cette variable :

- | | |
|--|--|
| Secteur DAF | Secteur OQN |
| <ul style="list-style-type: none"> • Inférieur à 80% • Entre 80% et 90% • 90% et plus | <ul style="list-style-type: none"> • Inférieur à 90% • Entre 90% et 100% • 100% et plus |

La mention spécialisée :

Il s'agit de la mention spécialisée du séjour/RHA spécifiée dans le PMSI. Trois modalités sont définies :

- Polyvalent : si le séjour n'a que des RHA polyvalents (aucune mention spécialisée)
- Spécialisé : si au moins un RHA du séjour a une mention spécialisée (quel que soit la spécialité)
- Gériatrique : si le séjour a au moins un RHA avec la mention « affections des personnes âgées polypathologiques, dépendantes ou à risque de dépendance », et aucun RHA avec une autre mention spécialisée

Le mode d'entrée :

Trois modalités sont définies :

- domicile,
- mutation
- transfert (définitif ou provisoire)

Le mode de sortie :

Quatre modalités sont définies :

- domicile,
- mutation
- transfert (définitif ou provisoire)
- décès

Le calage s'effectue indépendamment pour chacun des deux secteurs, GME par GME, pour les groupes présentant un nombre suffisant d'observations (au minimum 30 observations).

Chaque GME est donc calé non pas sur toutes ces variables, mais uniquement sur celles qui sont explicatives du coût de ces séjours/RHA.

3.2 Calculs de précision

L'ENC produit des estimations de coûts moyens par GME à partir d'un échantillon d'établissements – et donc d'un échantillon de journées. 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 d'établissements volontaires), de mesurer l'ampleur de cette sensibilité. Pour ce faire, la colonne intitulée « Erreur relative d'échantillonnage » de l'onglet « Caractéristiques statistiques » fournit l'erreur relative (ERE) de l'estimateur du coût moyen par GME. Cette erreur est due à l'aléa d'échantillonnage des établissements, après élimination des séjours/RHA extrêmes. Ce calcul d'erreur prend en compte la procédure de calage exposée dans la partie 3.1., dont l'objectif spécifique est de limiter l'incertitude liée à l'échantillonnage.

Depuis 2015, l'affichage des ERE dans le référentiel a été modifié pour faire apparaître des classes d'ERE :

- ERE ≤ 10% : **bonne** estimation de coût
- ERE entre 10% et 20% : estimation de coût **correcte**
- ERE ≥ 20% : estimation de coût **mauvaise**

Le calcul de l'ERE prend en compte les variables utilisées dans le calage au travers des résidus U_{ij} (où i correspond à l'établissement et j correspond au séjour ou à la journée) de la régression multiple des coûts C_{ij} sur l'ensemble des variables de calage. La régression est réalisée GME par GME. Pour un GME donné, la régression est effectuée sur l'ensemble des séjours/journées de l'échantillon de l'ENC (tous les types d'établissement apparaissent donc dans la même régression).

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

où : m_h est le nombre d'établissements du type h dans l'échantillon de l'ENC 2016,

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/journées par GME dans la base nationale 2016,

m_h est le nombre d'établissements du type h dans l'échantillon de l'ENC 2016 contribuant au GME,

M_h est le nombre d'établissements du type h dans la base nationale 2016 contribuant au GME.

Le nombre d'établissements de type CHU n'étant pas suffisant pour traiter ce type à part entière (il n'y a qu'un seul CHU dans l'échantillon 2016), les types CH et CHU ont été regroupés pour le 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 GME puisque les variables de calage sont propres à chaque GME.

Pour chaque GME, l'ERE est définie comme le rapport entre l'écart-type associé à l'erreur d'échantillonnage et le coût moyen estimé à partir des données 2016. Exprimée en pourcentage, l'ERE permet d'apprécier immédiatement l'incertitude relative qui entoure l'estimation du coût moyen du GME.

Le coût moyen réel – celui qui serait par définition obtenu à partir de l'ensemble des journées produites par l'ensemble des établissements du champ SSR – a environ 95 chances sur 100 d'être compris dans l'intervalle de confiance :

$$[\hat{C}(1 - 1,96.ERE); \hat{C}(1 + 1,96.ERE)]$$

Exemple :

Le coût moyen estimé pour un GME est 8 500 euros et l'ERE calculée vaut 5%. Il y a environ 95 chances sur 100 pour que le coût moyen réel soit compris entre 8 500 \times (1 - 1,96*5%) euros et 8 500 \times (1+1,96*5%) euros, soit un intervalle de confiance estimé (en euros) égal à [7 667 ; 9 333].

Pour certains GME il n'est pas possible d'estimer l'erreur d'échantillonnage (GME entièrement produit par un unique établissement ou GME produits par 2 établissements de 2 types différents). Dans ce cas, est mentionné NC (non calculé) dans la colonne concernée.

3.3 Indicateur de fiabilité

En complément de l'ERE, un indicateur de fiabilité de l'estimation du coût est calculé.

Cet indicateur est fondé sur le nombre de séjours (en HC) ou RHA (en HP), le nombre d'établissements produisant au moins 80% des séjours/RHA du GME, le taux de sondage et la précision de l'estimation (ERE).

3 scores sont définis : bon, correct et mauvais.

Le score est construit selon l'algorithme suivant :

- SCORE BON :
 - o Si le nombre de séjours (HC) ou RHA (HP) \geq 75 et nombre d'établissements \geq 3 et précision $<$ 10%
 - o Si le nombre de séjours (HC) ou RHA (HP) $<$ 75 et taux de sondage \geq 80%
- SCORE CORRECT :
 - o Si le nombre de séjours (HC) ou RHA (HP) $<$ 75 et taux de sondage entre 60% et 80% et précision $<$ 10%
- SCORE MAUVAIS :
 - o Si le nombre de séjours (HC) ou RHA (HP) $<$ 75 et taux de sondage $<$ 60% et [précision \geq 10% OU nombre d'établissements $<$ 3]

Les GME qui n'entrent dans aucune de ces modalités ont un score CORRECT.

Par ailleurs, les GME avec un score BON ou CORRECT sont dégradés en MAUVAIS s'ils sont dans une ou plusieurs de ces situations :

- une distorsion est observée sur le nombre de journées de présence ou sur l'âge entre l'ENC et la base nationale, sans que le calage ne la corrige
- les données de coûts ont été considérées comme insuffisamment fiables par rapport au contenu de la prise en charge médicale

Ces GME « dégradés » sont identifiables via un astérisque : leur indicateur de fiabilité vaut « MAUVAIS* ».