



CLASSIFICATION SSR TRAVAUX 2007-2008

**MODÉLISATION DES COÛTS EN SSR
ET CONSTRUCTION DE L'INDICATEUR DE
VALORISATION DE L'ACTIVITÉ (IVA)**

**AGENCE TECHNIQUE
DE
L'INFORMATION SUR L'HOSPITALISATION**

SOMMAIRE

1.	Introduction.....	4
2.	Contexte : préparer la T2A en SSR	4
3.	Les GHJ et leurs limites.....	4
4.	Méthodes.....	6
4.1	Données utilisées : ENC 2000-2001	6
4.2	Les outils statistiques.....	6
4.3	Choix de la méthode et étapes du traitement.....	6
4.3.1	Paradigme des coûts journaliers.....	7
4.3.2	D'une classification arborescente hiérarchique à une modélisation en modèle multivarié de régression linéaire.....	7
4.3.3	Les étapes de la modélisation.....	8
4.3.4	Les différents groupages étudiés.....	8
5.	Modifications du groupage : des GHJ aux GMD.....	8
5.1	Les principes du groupage en GMD	9
5.2	Evolution des listes	10
5.3	Hiérarchisation des tests	10
5.4	Variables prises en compte pour la détermination du GMD	10
5.5	Doubles tests.....	11
5.6	Suppressions et créations de tests, listes et GMD	11
5.7	Groupage en erreur	12
6.	La modélisation des coûts.....	13
6.1	Dictionnaire des variables utilisées.....	13
6.2	Analyse préliminaire uni et bivariée	15
6.3	Un modèle IVA _v 0.1. à plusieurs facteurs	16
6.4	Le modèle IVA _v 0.2.....	17
6.4.1	Dépendance cognitive	17
6.4.2	Les activités CdARR.....	18
6.4.3	La finalité principale	18
6.4.4	Les comorbidités et actes valorisants	19
6.4.5	Le principe des exclusions	20
6.5	Du modèle à son utilisation : la pondération en points IVA	21
6.5.1	La formule.....	21
6.5.2	Les classes de référence.....	22
6.5.3	Ajustement des coefficients.....	23
6.5.4	Les interactions.....	24
7.	Simulations et pertinence du modèle IVA	26

Sigles utilisés dans ce rapport

AE	Affection Etiologique
ARH	Agence Régionale de l'Hospitalisation
ATIH	Agence Technique de l'Information sur l'Hospitalisation
CAV	Comorbidités et Actes Valorisants
CCAM	Classification Commune des Actes Médicaux
CdAM	Catalogue des Actes Médicaux
CIM 10	Classification Internationale des Maladies, 10ème révision
CMC	Catégorie Majeure Clinique
CV	Coefficient de Variation
DAS	Diagnostic Associé Significatif
DG	Dotation Globale
DMS	Durée Moyenne de Séjour
DSS	Direction de la Sécurité Sociale
DRG	Diagnosis Related Groups
ENC	Étude Nationale de Coûts
ENCC	Étude Nationale de coûts à méthodologie commune
FG	Fonction Groupage
FHF	Fédération Hospitalière de France
FHP	Fédération de l'Hospitalisation Privée
FP	Finalité Principale de prise en charge
GHJ	Groupe Homogène de Journées
GHM	Groupe Homogène de Malades
GMD	Groupe de Morbidités Dominantes
ICR	Indice de Coût Relatif
IVA	Indicateur de Valorisation de l'Activité
MCO	Médecine Chirurgie Obstétrique
MMP	Manifestation Morbide Principale
MP	Morbidité Principale
MT2A	Mission Tarification à l'Activité
NGAP	Nomenclature Générale des Actes Professionnels
OQN	Objectif Quantifié National
PCL	Prise en charge Clinique Lourde
PCTL	Prise en charge Clinique Très Lourde
PMSI	Programme de Médicalisation des Systèmes d'Information
PVE	Part de Variance Expliquée
RF	Rééducation Fonctionnelle
RHA	Résumé Hebdomadaire Anonymisé
RSF	Résumé Standardisé de Facturation
RSFA	Résumé Standardisé de Facturation Anonyme
RHS	Résumé Hebdomadaire Standardisé
SSMed	Soins de suite médicalisés
SSR	Soins de Suite et Réadaptation
T2A	Tarification à l'activité

1. Introduction

Ce rapport présente les travaux statistiques - méthodes et résultats - réalisés pour mettre au point l'indicateur de valorisation de l'activité (IVA) et la nouvelle version du groupage en groupes de morbidités dominantes (GMD) qui l'accompagne, pour le PMSI en soins de suite et réadaptation (SSR), dont les principes sont décrits dans le *manuel de groupage SSR N°2009-1bis*⁽¹⁾.

Ces travaux se sont situés dans le cadre de la préparation à la tarification à l'activité en SSR.

Un rappel sur les limites de la classification antérieure en groupes homogènes de journées (GHJ) permet de comprendre les choix méthodologiques retenus, tels que : utilisation d'un modèle multivarié, simplification des groupes médicaux (passage des GHJ aux GMD), et analyse des autres variables contenues dans le résumé hebdomadaire standard (scores de dépendances, activités de rééducation-réadaptation, âge, type d'hospitalisation, comorbidités ou actes).

Les travaux ont été conduits en plusieurs étapes pour optimiser la manière d'utiliser chaque variable dans le modèle.

Les critères principaux pour juger de l'adéquation du modèle ont été la signification des paramètres utilisés et la part de variance expliquée des coûts journaliers.

2. Contexte : préparer la T2A en SSR

L'objectif fixé pour ces travaux était de proposer un modèle de valorisation de l'activité en SSR basé sur les données PMSI, afin de préparer la tarification à l'activité (T2A) dans ce champ.

Compte tenu des limites de la classification disponible en groupes homogènes de journées⁽²⁾ (GHJ), et du délai nécessaire pour disposer des données de la nouvelle étude nationale de coûts à méthodologie commune SSR (ENCc SSR) lancée par ailleurs, une phase de transition était nécessaire avant la mise en place d'une T2A en SSR.

L'enjeu était d'améliorer significativement la capacité à expliquer les coûts journaliers avec le recueil PMSI SSR, afin de disposer d'un outil tenant compte de l'activité médicale pour répartir les moyens en SSR. Cette classification en GMD et l'IVA ont été basés sur le principe en vigueur de la pondération des journées de présence et non pas sur un forfait au séjour, comme c'est le cas pour les groupes homogènes de malades (GHM) du champ MCO.

Il s'agissait également d'alimenter la réflexion et de progresser vers le modèle cible de valorisation et d'allocation des moyens de la T2A SSR.

3. Les GHJ et leurs limites

Compte tenu du polymorphisme du champ SSR en France, comprenant de la rééducation fonctionnelle, de la convalescence, de la gériatrie et des soins de suite médicalisés, la première classification SSR en GHJ⁽²⁾ mise au point en 1996 était une création purement française, construite essentiellement à partir d'une expertise médicale et d'un recueil limité de données financières, sans pouvoir s'appuyer sur une classification étrangère éprouvée, comme cela avait

été le cas pour le PMSI MCO, où les GHM avaient été bâtis sur la 4^{ème} version des Diagnosis Related Groups (DRG) américains.

Elle n'a fait l'objet d'aucune amélioration, en dehors de la création d'un GHJ spécifique pour les opérés cardiaques depuis moins de 21 jours dans la catégorie majeure clinique (CMC) « cardiovasculaire et respiratoire » (défini à dire d'experts à compter de 2001⁽³⁾).

L'étude nationale de coûts en SSR (ENC-SSR), relative à l'activité 2000 et 2001 d'une trentaine d'établissements du secteur sous dotation globale, avait permis de réaliser une pondération des GHJ, mais n'a pas été utilisée pour améliorer l'homogénéité des GHJ et leur pouvoir explicatif des coûts. Le rapport relatif à cette ENC-SSR⁽⁴⁾ avait mis en évidence une forte variabilité des coûts moyens d'un même GHJ d'un établissement à l'autre. Les établissements dits de « rééducation fonctionnelle » présentaient des coûts moyens par GHJ plus élevés que les établissements dits de « soins de suite médicalisée » d'environ 30%, la classification en GHJ prenant en compte de manière insuffisante les différences de prise en charge entre les établissements de rééducation fonctionnelle, de convalescence, de gériatrie ou autres suites médicalisées.

D'autres facteurs expliquaient également cette grande variabilité, parmi eux :

- une trop grande diversité des pathologies à l'intérieur d'un même GHJ avec une répartition différente entre les établissements. Si l'organisation des soins amène certains établissements à avoir majoritairement les cas lourds d'un GHJ alors que d'autres se concentrent plutôt sur les cas légers, alors les premiers sont forcément plus coûteux. La pathologie n'est pas le seul facteur de variation : la dépendance, l'âge, ou d'autres facteurs peuvent jouer un rôle. Par choix au moment de sa construction, la structure hiérarchique de la classification en GHJ a conduit à ne pas prendre en compte la dépendance dans de nombreux cas alors même qu'elle était identifiée comme une des variables à fort pouvoir explicatif des coûts ;
- la manière d'utiliser une variable. Dans un modèle de classification hiérarchique comme les GHJ, l'utilisation est dichotomique en « oui/non » et l'intégration d'une variable dans la classification ne signifie pas que les effets de cette variable sont complètement pris en compte. Par exemple les activités de rééducation-réadaptation n'étaient prises en compte que selon des bornes élevées et faisant intervenir au moins deux activités différentes : un patient bénéficiant, au cours d'une semaine de cinq jours, de quatre heures par jour d'une même et seule activité de rééducation-réadaptation était classé de la même manière (dans le même GHJ) qu'un autre patient pris en charge en convalescence sans aucune rééducation. De même les GHJ qui tenaient compte de la dépendance physique le faisaient par rapport à une borne du score de dépendance physique à 13 (\leq ou >12), les journées avec un score de dépendance à 4 étaient classées avec celles qui avaient un score de dépendance à 12.

Ces limites liées essentiellement aux principes des GHJ ne remettaient pas pour autant en cause la possibilité d'utiliser les données de coûts recueillies dans cette première étude nationale de coûts SSR, celle-ci comportant tous les principes fondamentaux des ENC du PMSI et permettant ainsi d'affecter une partie importante des coûts directement au patient, notamment :

- les dépenses de personnel infirmier et aide-soignant en fonction de son score SIIPS ;
- les dépenses de rééducation-réadaptation en fonction des minutes des activités de rééducation-réadaptation dont il avait bénéficié durant la semaine ;
- les dépenses relatives aux transports, produits sanguins labiles et actes réalisés à l'extérieur et facturés à l'établissement SSR ;
- les dépenses de prothèses, orthèses, implants et de médicaments coûteux.

La part des charges ventilées de manière forfaitaire à la journée était limitée essentiellement aux frais de logistique (1/3 des coûts, avec une très forte variabilité) et aux dépenses médicales, les coûts de structure étant exclus de l'échelle.

Les travaux suivants se sont donc appuyés sur ces éléments et constatations, avec pour objectif d'améliorer significativement la capacité à expliquer les coûts journaliers (part de variance expliquée des GHJ de 30% pour 261 groupes) en utilisant différemment les variables explicatives, déclarées dans le recueil PMSI SSR.

4. Méthodes

4.1 Données utilisées : ENC 2000-2001

Les résumés hebdomadaires anonymisés (RHA) utilisés pour cette étude, proviennent de la base de données de l'étude nationale de coûts SSR relative à l'activité 2000-2001 d'établissements volontaires du secteur sous dotation globale (sélectionnés après appel à candidature), soit 35 établissements au départ⁽⁴⁾.

Parmi ces établissements, quatre n'avaient finalement pas été intégrés dans la base ENC SSR 2000-2001 en raison d'un abandon pour deux d'entre eux et d'un défaut de qualité des données comptables pour les deux autres.

Les données exploitées ont donc été celles qui avaient fait l'objet des rapports ENC SSR 2000 puis 2000/2001 publiés respectivement en 2003 et 2004, sur le site de l'ATIH⁽⁴⁾, avec conservation des valeurs extrêmes (hors anomalies) décrites par les établissements. La base d'analyse ainsi constituée, comportait 316 696 RHA relatifs à l'activité de 31 établissements SSR.

4.2 Les outils statistiques

Les analyses ont été effectuées avec la requête de régression linéaire de SAS, sous modélisation standard, c'est-à-dire que les estimations ont été obtenues par maximisation du r^2 (qui correspond à la part de variance expliquée ou PVE) pour les coûts journaliers sans transformation (logarithmique ou autre) des variables.

Le premier principe était de rechercher le modèle qui permettait d'expliquer au mieux la variabilité des coûts journaliers tout en restant médicalement pertinent et acceptable.

Le second principe était un principe de parcimonie qui visait à limiter le nombre de degré de libertés (ddl) du modèle ou du moins d'optimiser le nombre de ddl pour une part de variance expliquée donnée.

4.3 Choix de la méthode et étapes du traitement

La construction de ce modèle intermédiaire a été guidée par les objectifs suivants, définis par la mission tarification à l'activité (MT2A) :

- soutenir l'évolution de l'offre de soins SSR et sa restructuration le cas échéant, pour inscrire les soins de suite dans une démarche de « cursus patient » sans rupture entre le court séjour, d'une part, le long séjour sanitaire et le médico-social d'autre part ;

- inciter à la fluidité des soins par un financement des activités SSR mieux adapté que le financement actuel ;
- prendre en compte l'augmentation des besoins liés au vieillissement de la population, au développement des maladies chroniques et du handicap ;
- mettre en place un système d'allocation de ressources plus équitable en fonction du poids de la prise en charge.

4.3.1 Paradigme des coûts journaliers

Compte tenu de la dispersion forte des durées de séjour observée par GHJ, nous sommes restés dans le paradigme antérieur de l'explication des coûts par jour de présence et non des coûts du séjour complet, comme en MCO ou dans certaines classifications étrangères propres à la rééducation fonctionnelle (ex Functional Related Groups).

Contrairement au PMSI MCO, il faut noter qu'il n'était pas possible de se référer directement aux classifications étrangères, le recueil et le périmètre des établissements étant très différents d'un pays à l'autre, ni même de s'affranchir des données de coûts pour construire la classification. En effet, les travaux sur la classification MCO en GHM s'appuient sur les bases nationales complètes en expliquant la part de variance des durées moyennes de séjours des GHM qui sont très fortement corrélées aux coûts.

En SSR, nous sommes dans le cadre des coûts journaliers, les tentatives d'explication des coûts par séjour sur la base de l'ENC SSR 2000/2001 n'ayant pas encore permis d'obtenir une part de variance expliquée suffisamment performante pour être utilisable dans un modèle de valorisation et d'allocation entre établissements.

La dépendance ressortait comme le principal facteur médical explicatif des coûts, mais la corrélation entre dépendance et coût n'était pas suffisante pour un modèle préparatoire à la T2A. De plus, construire un modèle classificatoire qui expliquerait la dépendance ne serait pas suffisant car elle est elle-même un élément qui conditionne la prise en charge mais pas toujours, en particulier en début de prise en charge en SSR, soit une résultante des moyens engagés pour réadapter le patient.

Enfin, compte-tenu des données à disposition et des délais impartis, l'objectif d'une refonte totale des listes de morbidités des GHJ, voire même de la classification SSR elle-même, n'a donc pas été retenu à ce stade.

4.3.2 D'une classification arborescente hiérarchique à une modélisation en modèle multivarié de régression linéaire

Les contraintes imposées (coûts journaliers, contenu du RHA en vigueur au moment de l'étude, données de coûts ENC 2000-2001, délais, PVE au séjour insuffisante) ont conduit à rechercher à optimiser l'utilisation des éléments disponibles et pour ce faire, à abandonner le modèle habituel de la classification arborescente hiérarchique du type des GHJ ou des GHM, au profit d'un modèle qui prenne en compte différentes variables explicatives des coûts de manière plus systématique.

Avec les GHJ, le classement dans un groupe de rééducation-réadaptation (PRR ou PRRC, cf. *Manuel des GHJ*⁽²⁾) exclu la prise en compte de la dépendance ; de même le classement dans un groupe de prise en charge clinique lourde (PCL ou PCTL, cf. *Manuel des GHJ*⁽²⁾) ne tient compte ni de la dépendance ni des activités de rééducation et réadaptation. Pour pouvoir améliorer

significativement la PVE l'effet du groupe médical a été complété par la prise en compte systématique de la dépendance, des activités de rééducation-réadaptation ainsi que d'autres variables a priori pertinentes. Les différentes variables ont donc été intégrées dans une modélisation multivariée qui a permis d'additionner leurs effets, aboutissant à la construction d'un indicateur de valorisation de l'activité (IVA).

4.3.3 Les étapes de la modélisation

La modélisation de l'IVA s'est déroulée en plusieurs étapes :

- analyse isolée de l'effet de chaque variable du recueil (selon différentes modalités d'utilisation dans le modèle) ;
- construction des GMD et étude de l'effet combiné bivarié avec le GMD et chacune des autres variables ;
- modélisation multivariée et optimisation du modèle IVA (version 0.1.) ;
- retour vers les experts (proposés par les fédérations d'établissements) pour validation et avis complémentaires ;
- travaux complémentaires relatifs aux suggestions des experts et amélioration du modèle IVA notamment sur les modalités à retenir pour certaines variables et les interactions dans le modèle (version 0.2.) ;
- analyse de la pertinence du modèle IVA v0.2. à partir des données établissements (base nationale SSR).

4.3.4 Les différents groupages étudiés

Les travaux se sont basés notamment sur trois versions de groupage SSR, appliquées sur l'échantillon utilisé (base ENC SSR 2000-2001), à savoir :

- le groupage en GHJ version 3.4. (dernière version officielle des GHJ, mise en œuvre en 2007),
- le groupage en GMD selon le modèle IVAv0.1., premier groupage en GMD, basé sur les listes de morbidités utilisées dans la fonction groupage en GHJ, tout en prenant en compte la liste et les caractéristiques des codes CIM10 en vigueur en 2007 (soit des modifications postérieures à celles en vigueur pour l'ENC SSR en 2001, relatives à l'utilisation de ces codes et à la création d'un certain nombre d'entre eux),
- le groupage en GMD selon le modèle IVAv0.2., version ayant permis de réaliser les simulations d'utilisation du modèle IVA dans l'allocation de ressources des établissements SSR.

5. Modifications du groupage : des GHJ aux GMD

La composante médicale de l'arbre de décision en GHJ basée sur le regroupement des morbidités en Catégories Majeures Cliniques, puis au sein d'une CMC en listes de morbidités dominantes (cf *Manuel des GHJ*⁽²⁾) a été conservée.

En revanche, les autres nœuds de la classification en GHJ, basés sur le test propre à la CMC 80-Gériatrie aiguë, sur la notion de prise en charge clinique lourde ou très lourde (PCL, PCTL), une prise en charge de rééducation-réadaptation (PRR, PRRC) ou fonction de la dépendance physique ≤ 12 , ont été supprimés au profit d'une prise en compte directe complémentaire de l'effet de chacune de ces variables dans l'IVA.

Le principe a donc consisté à ne conserver que les trois étapes suivantes de classement de l'arbre de décision en GHJ :

- la répartition en catégories majeures cliniques (CMC) ;
- la séparation pédiatrie / adulte, qualifiée de « sous-CMC », calée sur la borne des 18 ans (16 ans pour les GHJ) ;
- les grands groupes médicaux basés sur les morbidités, baptisés groupes de morbidités dominantes (GMD).

Ces modifications ont conduit à orienter les résumés uniquement vers des GMD, avec homogénéisation de la lisibilité médicale de l'activité.

5.1 Les principes du groupage en GMD

L'ensemble des tests effectués sur les informations du RHS constitue l'algorithme de groupage, définissant 69 groupes de morbidités dominantes (GMD) répartis en 12 catégories majeures cliniques (CMC). Il est appliqué à chaque RHS indépendamment des autres RHS constituant un même séjour hospitalier. Il s'appuie sur les informations suivantes : les codes de diagnostic, notamment la finalité de prise en charge, la manifestation morbide principale (et l'affection étiologique si besoin) et les diagnostics associés ainsi que l'âge. L'agence technique de l'information sur l'hospitalisation (ATIH) développe, maintient et diffuse un programme informatique exécutant cet algorithme, sous la dénomination « fonction groupage » (FG), programme source destiné à être intégré par les sociétés de service dans un système d'information hospitalier, ou celle de « programme de groupage » (PdG), programme à interface graphique.

Les catégories majeures cliniques (CMC) restent un niveau déterminant de classement des résumés hebdomadaires. Elles correspondent le plus souvent à un système fonctionnel.

Dans la majorité des cas, le déroulement de l'algorithme est le suivant :

le RHS est classé dans l'une des CMC selon le diagnostic que la fonction groupage retient comme morbidité principale du RHS, parmi la « Finalité de prise en charge », la « Manifestation morbide principale » et l'« Affection étiologique » ;

les GMD sont obtenus par un nombre variable de subdivisions des CMC à partir de l'âge, de la morbidité principale (décrite par les variables « Finalité de prise en charge », « Manifestation morbide principale » et « Affection étiologique » si besoin) et parfois des diagnostics associés significatifs.

L'algorithme de groupage en GMD utilise des listes de diagnostics et un arbre de décision, qui décrit l'ensemble des tests faits sur les informations du RHS pour le classer dans un GMD déterminé. C'est un arbre binaire, en oui et non, avec des nœuds et des fonctions de décision (cf *Manuel de groupage SSR 2009-1bis*⁽¹⁾).

La numérotation des GMD s'est basée sur les numéros des listes initiales de la classification en GHJ (Ex : GMD « Tétraplégies âge sup. à 18 ans » de la CMC 12, correspondant à la liste 1201 a le numéro 12A01 (12= CMC, A=sous-CMC Adultes, 01=2 derniers chiffres du n° de la liste initiale).

En raison d'effectifs très faibles (<0,5% dans l'échantillon étudié et la base nationale SSR 2007) et afin d'améliorer la lecture médicale du case-mix, la CMC « 80-Gériatrie aigue » n'a pas été conservée. Il en a été de même pour la CMC « 30-attente de placement », qui était de plus soumise aux choix des établissements en matière d'affichage (mise en avant ou non de ces situations). Les résumés concernés ont donc été orientés vers les GMD correspondant aux morbidités principales décrites.

Plusieurs types de modifications ont été apportés dans l'arbre, et ont concerné :

- le contenu des listes de morbidités dominantes,
- la hiérarchisation dans l'arbre de décision des groupes de morbidités dominantes (ordre des tests dans chaque sous-CMC),
- les variables prises en compte (morbidité principale et DAS)
- les tests prenant en compte deux listes de morbidités dominantes,
- la création et la suppression de groupes de morbidités dominantes par rapport aux listes initiales de la classification en GHJ,
- le groupe erreur

5.2 Evolution des listes

Les premiers travaux se sont basés sur les listes de morbidités dominantes de la classification en GHJ en vigueur en 2007 (version 3.4 de la fonction groupage SSR) ; puis un certain nombre de modifications ont été apportées en s'appuyant sur l'analyse des données de coût et à dire d'experts (listes concernées : n°1402, 1701, 1702, 5003, 5004 et 5007), afin d'améliorer leur cohérence médicale.

Pour rendre l'arbre plus explicite et contrôlable, toutes les listes, y compris celles correspondant au dernier GMD de chaque sous-CMC (par défaut), ont été construites ; leur contenu est détaillé dans le *Manuel de groupage PMSI SSR Bulletin Officiel N°2009-1bis⁽¹⁾*.

5.3 Hiérarchisation des tests

La hiérarchie des GMD dans une sous-CMC (ordre des tests de la fonction groupage) s'est basée sur leur coût moyen journalier respectif (ordre décroissant), excepté pour les GMD de type « Autres ... » positionnés systématiquement en fin d'arbre quelque soit leur coût. Par rapport à la hiérarchisation initiale des listes de morbidités dominantes dans la classification en GHJ, quatre groupes ont été modifiés et positionnés en tête de l'arbre dans leur sous-CMC :

- GMD 11A01 « Insuffisances cardiaques, insuffisances respiratoires chroniques, âge supérieur à 18 ans »,
- GMD 12E03 « Déficiences cérébrales traumatiques, âge inférieur à 18 ans » et GMD 12A03 « Déficiences cérébrales traumatiques, âge supérieur à 18 ans »,
- GMD 16A03 « Pathologies rachidiennes, âge supérieur à 18 ans »,
- GMD 17A05 « Brûlures et corrosions, âge supérieur à 18 ans »

5.4 Variables prises en compte pour la détermination du GMD

Rappel : l'orientation dans une CMC ne peut-être déterminée que par l'un des trois codes de morbidité principale (FPP, MMP, AE), les DAS ne jouant aucun rôle à cette étape initiale. Cette règle est restée inchangée, puisqu'elle existait précédemment dans le groupage en GHJ.

A l'étape des travaux relatifs au groupage en GMD selon le modèle IVA0.1., les tests réalisés pour déterminer le GMD étaient effectués sur l'ensemble des codes CIM10 renseignés dans le

recueil, soit la morbidité principale et les diagnostics associés significatifs ; le GMD pouvait donc être déterminé à partir d'un DAS.

C'est à compter de la version v0.2. des GMD que le rôle classant des DAS a été limité et que seuls les trois diagnostics décrivant la morbidité principale ont été retenus pour définir l'orientation du recueil dans un GMD donné (exception faite des GMD basés sur un double test, décrit ci-dessous).

5.5 Doubles tests

Quelques GHJ des CMC 17 et 50 étaient déterminés par la présence d'au moins deux diagnostics dans une même liste, ou deux diagnostics appartenant chacun à deux listes distinctes de morbidités dominantes, quelque soit leur position (double test).

Dans la version v0.1., une contrainte supplémentaire a été introduite relative au caractère « distincts » des deux diagnostics recherchés ; contrairement au groupage en GHJ qui permettait de classer un recueil dans un GHJ « compliqué » lorsqu'il contenait deux fois un même code CIM10 (qui appartenait aux deux listes testées).

De fait, dans la version v0.2., ces doubles tests ont donc été basés sur les principes suivants :

- présence de deux codes CIM10 qui doivent différer sur leurs 3 premiers caractères et
- un des deux codes CIM10 doit obligatoirement être renseigné en morbidité principale (FPP, MMP, AE).

5.6 Suppressions et créations de tests, listes et GMD

Suppression des tests suivants, initialement présents dans la classification en GHJ :

- « 80 ans, FP=ZAIGU et MP=diagnostic d'entrée de la CMC 80 » car suppression de la CMC 80-Gériatrie aiguë,
- « soins post-chirurgicaux » à partir de la date de la dernière intervention chirurgicale dans les sous-CMC11-Pédiatrie, sous-CMC11-Adulte, sous-CMC15-Pédiatrie, sous-CMC15-Adulte et sous-CMC16-Adulte,
- « avoir au moins deux diagnostics de la liste n°1703 » pour le GMD 17A03 dans la sous-CMC17-Adulte,
- « avoir au moins un diagnostic de la liste n°1801 » sans aucune autre condition pour le GMD 18A01 dans la sous-CMC18-Adulte,
- « avoir au moins un diagnostic de la liste n°2001 », car suppression du groupe spécifique aux maladies par VIH dans la sous-CMC20-Adulte,
- « avoir au moins un diagnostic de la liste n°3001 », « avoir au moins un diagnostic de la liste n°3002 » et « avoir au moins un diagnostic de la liste n°3003 » car suppression de la CMC30,
- « avoir au moins un diagnostic de la liste n°4001 » et « avoir au moins un diagnostic de la liste n°4002 » car un seul GMD retenu, le 40A22 dans la sous-CMC40-Adulte

Création de nouvelles listes de morbidités dominantes, telles que :

- liste n°1405 car différente de la liste initiale n°1404 dans la sous-CMC14-Pédiatrie pour le GMD 14E05,
- liste n°1501 afin de regrouper les hémopathies et certaines tumeurs malignes (classées pour certaines initialement dans la CMC60-Autres situations) dans les GMD15E01 et GMD15A01 de la CMC15-Hémopathies malignes et affections des organes digestifs et génitaux urinaires,

- liste n°1803 contenant les amputations de l'épaule, du bras, de la hanche et les amputations multiples, pour le test « avoir au moins un diagnostic dans la liste n°1803 » qui oriente vers le GMD 18A01 de la sous-CMC18-Adulte,
- liste n°1804 contenant les autres amputations, pour le double test « avoir au moins deux diagnostics dans la liste n°1804 » qui oriente vers le GMD 18A01 de la sous-CMC18-Adulte,
- liste n°1802 contenant les autres amputations et ajustement de prothèse externe, pour le double test « avoir au moins un diagnostic dans la liste n°1802 et avoir au moins un diagnostic dans la liste n°1801 », qui oriente vers le GMD 18A01 de la sous-CMC18-Adulte

5.7 Groupage en erreur

Pour ces travaux, deux groupes erreur ont été isolés :

- GMD 9000B-Non groupable : quand au moins un code CIM10 est renseigné dans une position non autorisée ou lorsque les codes CIM10 renseignés en position principale (FP, MP, AE) ne permettent pas de classer le recueil dans une CMC ou un GMD donné,
- GMD 9000A-Autres erreurs : lorsque le recueil est reconnu comme un doublon ou dont le nombre de jour de présence est égal à zéro.

Toutes ces modifications ont donc aboutit à la création de 42 GMD Adultes et 27 GMD Enfants suivants :

CMC	N° GMD ADULTES	Libellé GMD
11	11A01	INSUF. CARD. ET INSUF. RESP. CHRON., AGE SUP. A 18 ANS
11	11A02	T.M. DE L'APP. RESP. ET AFF. RESP. LIEES AU VIH, AGE SUP. A 18 ANS
11	11A20	AUTRES AFF. DE L'APP. CIRC. ET DE L'APP. RESP., AGE SUP. A 18 ANS
12	12A01	TETRAPLEGIES, AGE SUP. A 18 ANS
12	12A02	LES. CEREB. NON TRAUMA., AGE SUP. A 18 ANS
12	12A03	LES. CEREB. TRAUMA., AGE SUP. A 18 ANS
12	12A04	PARAPLEGIES, AGE SUP. A 18 ANS
12	12A05	MONOPLÉGIES, AGE SUP. A 18 ANS
12	12A22	AUTRES AFF. DU SYST. NERV., AGE SUP. A 18 ANS
13	13A01	DEMENCE, SCHIZOPHRENIE ET RETARD MENT., AGE SUP. A 18 ANS
13	13A02	DEPEND. ET Tbl. MENT. ORG. DUS A SUBST. PSYCHOACT., AGE SUP. A 18 ANS
13	13A03	TROUBLES DEPRESSIFS ET ANXIEUX, AGE SUP. A 18 ANS
13	13A20	AUTRES Tbl. MENT., AGE SUP. A 18 ANS
14	14A01	ULCERES DE DECUBITUS, AGE SUP. A 18 ANS
14	14A02	CECITE ET DEFIC. DE LA VISION, AGE SUP. A 18 ANS
14	14A04	AUTRES AFF. DE L'OEIL, DE L'OREILLE ET CERTAINS Tbl. SENS., AGE SUP. A 18 ANS
14	14A20	AUTRES AFF. DE LA PEAU ET DES TISSUS SOUSCUT., AGE SUP. A 18 ANS
15	15A01	HEMOPATHIES ET CERTAINES TUMEURS MALIGNES, AGE SUP. A 18 ANS
15	15A22	AFF. DES ORG. DIG ET GEN-URIN. Sf. T.M., AGE SUP. A 18 ANS
16	16A01	ARTHROPAT. NON INFECT., AGE SUP. A 18 ANS
16	16A02	ARTHROPAT. INFECT. ET INFL., OSTEITES ET MYOPATHIES, AGE SUP. A 18 ANS
16	16A03	AFF. DU RACHIS, AGE SUP. A 18 ANS
16	16A04	MALF., DEFORM. ET DYSTROPHIES OS-ART., AGE SUP. A 18 ANS
16	16A05	COMPLIQ. ET SUIVI D'IMPLANTS OS-ART., AGE SUP. A 18 ANS
16	16A20	AFF.OS-ART. NCA, AGE SUP. A 18 ANS
17	17A01	LES. OS-ART. ET TENDONS TRAUMA., Sf POIGNET MAIN, AGE SUP. A 18 ANS
17	17A03	LES. TRAUMA. MULT., COMPLEXES OU COMPLIQ., AGE SUP. A 18 ANS
17	17A04	LES. OS-ART. ET TENDONS TRAUMA., POIGNET MAIN, AGE SUP. A 18 ANS
17	17A05	BRULURES ET CORROSIONS, AGE SUP. A 18 ANS
17	17A08	FRAC. DU COL DU FEMUR AVEC POSE DE PROTH. DE HANCHE, AGE SUP. A 18 ANS

17	17A09	FRAC. DU COL DU FEMUR AVEC OS., AGE SUP. A 18 ANS
17	17A20	AUTRES LES. DU SYST. OS-ART. TRAUMA., AGE SUP. A 18 ANS
18	18A01	AMPUT. EPAULE, BRAS, HANCHE, MULT. ET AVEC CMA, AGE SUP. A 18 ANS
18	18A20	AUTRES AMPUTATIONS, AGE SUP. A 18 ANS
20	20A22	SOINS PALLIATIFS, AGE SUP. A 18 ANS
40	40A22	READAPTATION ET REINSERTION, AGE SUP. A 18 ANS
50	50A02	DIABETE COMPLIQ., AGE SUP. A 18 ANS
50	50A03	DIABETE ET OBESITE AVEC MAL. ASSOC., AGE SUP. A 18 ANS
50	50A05	OBESITE SANS MAL. ASSOC., AGE SUP. A 18 ANS
50	50A06	DENUTRITIONS MAJEURES, AGE SUP. A 18 ANS
50	50A20	AUTRES AFF. NUTR., AGE SUP. A 18 ANS
60	60A22	AUTRES AFF., AGE SUP. A 18 ANS

CMC	N° GMD ENFANTS	Libellé GMD
11	11E03	MUCOVISCIDOSE
11	11E21	AUTRES AFF. DE L'APP. CIRC. ET DE L'APP. RESP., AGE INF. A 18 ANS
12	12E01	TETRAPLEGIES, AGE INF. A 18 ANS
12	12E02	LES. CEREB. NON TRAUMA., AGE INF. A 18 ANS
12	12E03	LES. CEREB. TRAUMA., AGE INF. A 18 ANS
12	12E04	PARAPLEGIES, AGE INF. A 18 ANS
12	12E05	MONOPLÉGIES, AGE INF. A 18 ANS
12	12E22	AUTRES AFF. DU SYST. NERV., AGE INF. A 18 ANS
13	13E04	RETARD MENT. ET CERTAINS Tbl. DU DEVELOP., AGE INF. A 18 ANS
13	13E21	AUTRES Tbl. MENT. ET DU COMPORT., AGE INF. A 18 ANS
14	14E05	AFF. DE L'OEIL, DE L'OREILLE ET Tbl. SENS., AGE INF. A 18 ANS
14	14E21	AFF. DE LA PEAU ET DES TISSUS SOUSCUT., AGE INF. A 18 ANS
15	15E01	HEMOPATHIES ET CERTAINES TUMEURS MALIGNES, AGE INF. A 18 ANS
15	15E22	AFF. DES ORG. DIG ET GEN-URIN. Sf. T.M., AGE INF. A 18 ANS
16	16E04	MALF., DEFORM. ET DYSTROPHIES OS-ART., AGE INF. A 18 ANS
16	16E21	AFF. RHUMATO. DU SYST. OS-ART. NCA, AGE INF. A 18 ANS
17	17E01	LES. OS-ART. ET TENDONS TRAUMA., Sf POIGNET MAIN, AGE inf. A 18 ANS
17	17E21	LES. TRAUMA. NCA, AGE INF. A 18 ANS
18	18E21	AMPUTATIONS, AGE INF. A 18 ANS
20	20E22	SOINS PALLIATIFS, AGE INF. A 18 ANS
40	40E22	READAPTATION ET REINSERTION, AGE INF. A 18 ANS
50	50E01	ANOM. METAB. ET CERTAINES MAL. DE L'APP. DIG., AGE INF. A 18 ANS
50	50E07	OBESITE, AGE INF. A 18 ANS
50	50E08	D.I.D., AGE INF. A 18 ANS
50	50E09	ANOREXIE MENT., AGE INF. A 18 ANS
50	50E21	D.N.I.D. ET AUTRES AFF. NUTR., AGE INF. A 18 ANS
60	60E22	AUTRES AFF., AGE INF. A 18 ANS

6. La modélisation des coûts

6.1 Dictionnaire des variables utilisées

Soit l'ensemble des variables qui ont été testées dans la modélisation, en plus du GMD, classées par famille :

- **Sous-CMC** : correspond aux 2 branches distinctes qui existent pour chaque CMC entre les GMD adultes et les GMD enfants. C'est un des niveaux de stratification des coefficients IVA utilisé pour tenir compte des interactions entre les variables du modèle IVA

- **PED** : séparation transversale selon l'âge (moins de 18 ans, 18 ans ou plus), appliquée en amont des CMC. C'est un des niveaux de stratification des coefficients IVA utilisé pour tenir compte des interactions entre les facteurs
- **AGE** : utilisé en classes (variable qualitative ordinale) avec les classes suivantes :
 - pour la version v0.1. : 0-6 / 7-17 / 18-30 / 31-60 / 61-75 / >75
 - pour la version v0.2. : 0-5 / 6-12 / 13-17 / 18-29 / 30-59 / 60-74 / >74
- **Dep-Phys** : score des 4 items de dépendance physique (alimentation, habillage, continence et déplacement), valeurs de 4 à 16
- **Dep-Cogn** : score des 2 items de dépendance cognitive (comportement et communication), valeurs de 2 à 8
- **Dep-Tot** : scores de dépendances sur les 6 items de dépendance, valeurs de 6 à 24
- **TypHosp** : type d'hospitalisation, en deux classes : hospitalisation complète ou temps partiel (hospitalisation de jour, de nuit ou séances, traitements et cures ambulatoires)
- **RF** : discipline de l'établissement déclarée dans le cadre de l'ENC, avec 3 classes : RF (rééducation fonctionnelle), SSMéd (soins de suite médicalisés) et mixte.
Dans la version v0.2., cette variable a été remplacée par les informations renseignées dans « Finalité principale de prise en charge » (Finalp) car cette discipline n'était pas déclarée dans le format du RHA
- **Plusieurs variantes ont été testées pour la variable Finalp** (parmi toutes les modalités de finalité principale de prise en charge) :
 - **Finalp1** : Finalité principale de prise en charge décrite avec un code CIM10 commençant par Z44 ou Z50 (toutes les autres modalités étant non significatives vis-à-vis du coût journalier)
 - **Finalp2** qui comprend les modalités suivantes :
 - « Z50.1-Autres thérapies physiques » en FP
 - « Z50.8-Soins impliquant d'autres moyens de rééducation » en FP
 - « Z50.0-Rééducation cardiaque » en FP et regroupement du RHS dans la CMC11
 - « Z44.0-Mise en place et ajustement d'un bras artificiel » ou « Z44.1-Mise en place et ajustement d'une jambe artificielle » en FP et le regroupement du RHS dans la CMC18
- **Plusieurs variantes ont été testées pour les activités de rééducation** (parmi les 12 décrites dans le RHA) :
 - **DureeReed** : durée totale des temps de rééducation-réadaptation
 - **NbReed1-8** : nombre d'activités de rééducation-réadaptation avec au moins une minute indiquée dans le RHA. NbReed1-8 ne prend en compte que les 8 activités individuelles classantes en GHJ, excluant les activités « bilan », « physiothérapie », « balnéothérapie » et « rééducation collective »
 - **NbReed30-8** : nombre d'activité de rééducation (toujours parmi les 8 classantes) dont le temps est ≥ 30 min/jour
 - **NbReed1-12** : nombre d'activité de rééducation-réadaptation, parmi les 12, avec au moins une minute déclarée, sachant que la valeur 1 est considérée comme la valeur zéro (si $\text{NbReed1-12} \leq 1$ alors $\text{NbReed1-12} = 0$)
 - **NbReed5-12** : nombre d'activité de rééducation-réadaptation, parmi les 12, ayant au moins 5 minutes par jour de présence. Dans ce décompte, les 4 activités supplémentaires (rééducation collective, bilan, physiothérapie et balnéothérapie) ne comptent que pour 0,5 chacune au lieu de 1 dans le score.
- **PCTL** : présence d'un acte ou d'un diagnostic de la liste des PCTL (cf *Manuel des GHJ*⁽²⁾)

- **PCL** : présence d'un acte ou d'un diagnostic de la liste des PCL, mais avec absence de PCTL (cf *Manuel des GHJ*⁽²⁾)
- **CAV** : présence d'un DAS ou d'un acte CCAM appartenant à la liste des CAV (comorbidités et actes valorisants, cf *Manuel de groupage PMSI-SSR n°2009-1bis*⁽¹⁾)
- **ZAIGU** : présence du code « ZAIGU » en finalité principale (cf *Guide de production des RHS*⁽⁵⁾)
- **ZAIGU80** : présence du code « ZAIGU » en finalité principale et âge >79 ans (RHA classé auparavant en CMC 80-Gériatrie aiguë, cf *Manuel des GHJ*⁽²⁾)
- **Fauteuil** : « utilisation d'un fauteuil roulant = Oui » dans le RHA (cf *Guide de production des RHS*⁽⁵⁾)
- **Chir** : nombre de jours après l'intervention chirurgicale déclaré dans le RHA (cf *Guide de production des RHS*⁽⁵⁾), utilisé en 4 classes : sans chirurgie / 0-20 / 21-90 / >90 jours

Les tests ont donc porté sur l'ensemble des variables utilisées précédemment par le groupage en GHJ ainsi que quelques autres.

6.2 Analyse préliminaire uni et bivariée

La première analyse a déterminé les parts de variance expliquées (PVE) de chacun des facteurs, seuls ou en association avec les CMC, les sous-CMC ou les GMD. C'est-à-dire que, pour chaque facteur, F (par exemple Dep-Phys), 4 types d'association ont été établis :

- aucune association (uniquement F)
- F et CMC
- F et sous-CMC
- F et GMD

Le nombre de paramètres indépendants (ou degré de liberté noté ddl) dans chaque modèle est égal au nombre de paramètres pour le facteur F (1 pour une variable utilisée en continu comme les scores de dépendance, et [nombre de classes -1] pour les facteurs en classes) multiplié par le nombre de CMC, sous-CMC ou GMD.

Tableau A : PVE par facteur et association

Facteur	Ddl	PVE (%) en association avec			
		Aucune	CMC	SousCMC	GMD
			10.15	13.21	18.82
DEP-Cogn	1	7.82	18.13	20.54	26.41
DEP-Phys	1	23.38	27.69	29.74	33.08
DEP-Tot	1	21.77	27.76	29.80	33.34
Age	5	2.54	15.08	15.08	20.63
PCTL+PCL	2	5.42	14.78	17.52	21.79
ZAIGU	1	0.001	10.45	13.59	19.24
ZAIGU80	1	0.03	10.33	13.44	19.05
CHIR	3	0.19	10.87	14.21	20.28
TypHosp	1	0.50	12.03	15.67	21.40
NbReed1-8	1	12.80	19.28	21.87	26.62
NbReed30-8	1	5.93	14.20	17.33	22.74
NbReed1-12	1	12.91	19.16	21.75	26.51
Fauteuil	1	9.90	15.35	17.92	22.69
Finalp1	1	2.27	11.74	15.35	20.73
RF	2	12.36	19.90	22.53	27.23

Les GMD à eux seuls aboutissaient à une PVE de 18.82%, soit un peu moins que la dépendance totale seule.

La dépendance physique est apparue comme le principal facteur d'explication des coûts avec 23% de PVE. Compte tenu des interactions entre ces 2 variables, le modèle bivarié combinant la dépendance avec les GMD (modèle de régression à 2 variables) n'est pas simplement l'addition de l'effet isolé des variables, avec une PVE de 33,08% (déjà supérieur à la PVE du modèle en GHJ (29.85%).

Le groupe de facteurs suivant était relatif au nombre d'activités de rééducation-réadaptation (NbReed), vient ensuite la discipline (RF) ; les autres variables n'apportant que peu d'informations en plus des GMD.

Un test statistique comparant les deux modèles : [sousCMC + sousCMC * (DEP-Phys)] et [GMD + GMD*Dep-Phys] montre que l'augmentation de la PVE est statistiquement significative ($p < 0.0001$) (29.74% contre 33.08%, avec passage de 50 à 146 ddl) avec l'ajout d'une variable au GMD, et montre également que le coefficient de DEP-Phys n'est pas constant par sous-CMC. On considérera donc par la suite qu'il faut bien descendre au niveau du GMD pour constituer les groupes de base.

Il en résulte que la base du modèle devait prendre les GMD et la dépendance physique, avec interaction entre les deux. La base du modèle comporte une équation de type [IVA = A x Dep-Phys + B] par GMD). On été ajouté ensuite à ce modèle de base, les facteurs nécessaires, incluant si possible des interactions avec la sous-CMC ou interaction avec seulement le facteur PED, pour limiter le nombre de paramètres du modèle.

6.3 Un modèle IVAv0.1. à plusieurs facteurs

Un premier modèle multivarié a été testé, contraint de telle sorte que certaines variables ne dépendaient plus du GMD. Ce choix a été établi sur des validations statistiques et pour restreindre le nombre de coefficients à estimer.

Le modèle initial était de la forme :

$$\text{IVA (PED=p, GMD=g, Sous-CMC=s, DepPhy=dp, NbReed1-8=r, Age=a, Finalp1=f, TypHosp=t)} \\ = \\ \text{GMD(g) + DepPhy(g) * dp + NbReed1-12(s) * r + Age(s,a) + Finalp1(s,f) + TypHosp(p,t)}$$

Le nombre de points IVA journalier moyen pour un patient dans le GMD n°g, pour la sous CMCs, pour une séparation adultes/enfants=p, et ayant pour les autres facteurs les valeurs suivantes : Dépendance Physique=dp, Nombre d'activité Cdarr=r, âge dans la classe n°a, une finalité de prise en charge=f et un type d'hospitalisation=t sera la somme de 6 termes.

Le coefficient de la « dépendance physique » varie d'un GMD à l'autre ; les coefficients de la variable « finalp1 », les « classes d'âge » et le coefficient du « nombre d'activités de rééducation-réadaptation » ont été estimés au niveau de la sous-CMC ; les coefficients de la variable « type d'hospitalisation » ont été estimés au niveau de la distinction adultes/enfants.

La part de variance correspondante est de 42.35% (pour 249 ddl).

6.4 Le modèle IVAv0.2.

Un certain nombre de propositions d'amélioration ont été définies par un groupe de travail composé d'experts (désignés par les fédérations), telles que :

- la dépendance cognitive utilisée en tant que variable supplémentaire,
- la suppression de la CMC30 avec réaffectation des RHS dans les autres CMC,
- la création d'un seul GMD Adultes pour la CMC40,
- la prise en compte des 12 activités CdARR (au lieu de 8) avec un seuil de temps minimum de 5mn par journée de présence,
- la création d'un GMD 15A01 et 15E01 "Hémopathies et certaines tumeurs malignes" dans la CMC 15,
- la prise en compte plutôt en classes de certaines modalités de « finalité principale » (cf détail dans chapitre 6.1 - dictionnaire des variables utilisées),
- l'introduction d'une variable supplémentaire dans le modèle prenant en compte certaines comorbidités (liste fermée de codes CIM10 et CCAM),
- l'utilisation du principe d'exclusion pour les CAV,
- des améliorations relatives aux listes de diagnostics d'entrée des CMC,
- des améliorations relatives aux listes de diagnostics des GMD,
- des évolutions concernant l'arbre de groupage (ajouts et suppressions d'embranchements cf. chapitre 5.2).

6.4.1 Dépendance cognitive

Dans la version v0.1., cette variable n'avait pas été intégrée, son absence de rôle dans la classification en GHJ, ayant incité à une grande prudence dans son utilisation. La première analyse des PVE (cf. tableau A) sur les facteurs pris séparément, montrait pour un score de dépendance calculé à partir des 6 variables (habillement, déplacement, alimentation, continence, comportement et communication) une PVE plus petite que pour un score calculé sur les quatre variables dites de dépendance physique (habillement, déplacement, alimentation et continence avec 21.7 vs 23.3). De plus, le score de dépendance cognitive (comportement et communication) pris sans aucune interaction faisait partie des facteurs avec un effet faible.

Suite aux propositions d'améliorations faites par les groupes d'expert, la possibilité d'intégrer la dépendance cognitive dans le modèle a été revue en partant cette fois du modèle v0.1. (et non pas de la base brute de l'ENC-SSR).

Le premier test s'est porté sur la dépendance cognitive en interaction avec le GMD (comme pour la dépendance physique) avec une augmentation de 1,5 % pour 73 Ddl de la PVE.

Ces résultats ont permis de confirmer qu'il était pertinent de prendre en compte la dépendance cognitive dans l'IVA et qu'en l'absence d'une corrélation positive entre les coefficients de dépendance physique et cognitive, il était préférable de l'intégrer en tant que variable supplémentaire plutôt qu'en additionnant le score de dépendance physique avec celui de la dépendance cognitive.

6.4.2 Les activités CdARR

Pour chaque RHA, douze activités de rééducation-réadaptation sont identifiées (et définies chacune par une liste fermée d'actes dans le Catalogue des Activités de Rééducation-Réadaptation CdARR), dont trois d'entre elles : « Bilan », « Physiothérapie » et « Balnéothérapie » avaient été jugées non discriminantes dans la classification en GHJ, ce qui avait conduit à ne retenir que les huit activités individuelles de rééducation-réadaptation (classantes en GHJ, cf *Manuel des GHJ*⁽²⁾) dans le premier modèle v0.1. pour construire la variable « NbReed » (voir définition dans le chapitre 6.1).

A la demande du groupe d'expert, la possibilité d'intégrer les trois autres activités individuelles et l'activité « rééducation collective » dans le modèle a été étudiée, ainsi que la définition d'un seuil minimum de temps journalier.

Ce choix d'un seuil minimal de minutes par jour pour prendre en compte une activité donnée a été établi sur un critère de recherche de maximisation de la part de variance expliquée du modèle v0.1. Sur les six seuils testés suivants : au moins une minute, au moins 5 minutes, au moins 10 minutes, au moins 20 minutes, au moins 30, le seuil retenu maximisant la PVE a été d'au moins 5 minutes par jour de présence (NbReed5-12).

Tableau B : PVE pour « Nombre d'activité CdARR »

	Temps moyen >0 min/jour		Temps moyen >=5 min/jour	
	8 items (NbReed1-8)	12 items (NbReed1-12)	8 items (NbReed5-8)	12 items (NbReed5-12)
NbReed	41.62	42.22	42.44	42.79

Afin d'atténuer les effets modestes de ces quatre activités complémentaires, un coefficient à 0,5 a été utilisé dans la construction de la variable « NbReed5-12 » permettant ainsi de pondérer l'effet de leur introduction (voir détails dans le *manuel de groupage SSR 2009-1bis*⁽¹⁾).

6.4.3 La finalité principale

Ne disposant pas dans les informations contenues dans le RHS, d'une variable identifiant clairement les unités de type rééducation réadaptation fonctionnelle et celles de soins de suite médicalisés (facteur le plus explicatif des coûts journaliers dans l'ENC 2000-2001), une variable complémentaire du « nombre d'activité CdARR » (décrite dans le chapitre précédent) a été définie (appelée « finalp1 » dans le modèle V0.1.) pour approcher cet effet discriminant dans la consommation de ressources des patients.

Cependant, dans le cadre des premières simulations du modèle v0.1. de l'IVA avec les trois régions pilotes, l'utilisation globale des catégories de diagnostics CIM10 Z44 et Z50 (cf définition de la variable « finalp1 » dans le chapitre 6.1.) s'est avérée insatisfaisante.

Dans le modèle V0.2. les codes suivants n'ont donc pas été retenus : Z50.2-Sevrage d'alcool, Z50.3-Rééducation des drogués et après abus de médicaments, Z50.4-Psychothérapie, non classée ailleurs, Z50.5-Rééducation du langage, Z50.6-Education de la vision, Z50.7-Ergothérapie et rééducation professionnelle, non classées ailleurs, Z50.9-Soin impliquant une rééducation, sans précision, Z44.2-Mise en place et ajustement d'un œil artificiel, Z44.3-Mise en place et ajustement d'une prothèse externe du sein, Z44.8-Mise en place et ajustement d'autres appareils de prothèse externe, Z44.9-Mise en place et ajustement d'un appareil de prothèse externe non précisé, car soit ils ne caractérisaient pas des prises en charge de rééducation et réadaptation fonctionnelle, soit ils n'étaient pas assez précis, soit les effectifs en RHA étaient trop faibles.

Seuls les codes suivants ont été conservés pour constituer la variable « finalp2 », avec pour certains une contrainte supplémentaire liée à leur caractère classant dans une CMC donnée.

Tableau C : Modalités de la FP pour la variable « finalp2 »

Code CIM10 en Finalité Principale de prise en charge	libellé	Classant dans CMC
Z44.0	Mise en place et ajustement d'un bras artificiel	18
Z44.1	Mise en place et ajustement d'une jambe artificielle	18
Z50.0	Rééducation des cardiaques	11
Z50.1	Autres thérapies physiques	99
Z50.8	Soins impliquant une rééducation, sans précision	99

Une étape de validation statistique et médicale des coefficients estimés a été ensuite réalisée, ayant conduit à ne retenir les modalités « Z44.0 », « Z44.1 » et « Z50.0 » que dans certaines sous-CMC du modèle (cf définition de la variable « finalp2 » dans le chapitre 6.1.).

Les coefficients de ces facteurs ont été estimés au niveau sous-CMC.

6.4.4 Les comorbidités et actes valorisants

L'introduction de ce facteur a eu pour but de remplacer les GHJ de type PCL et PCTL qui concernaient les séjours avec une prise en charge particulièrement lourde, recherchée sur une liste de diagnostics et d'actes médicaux (il existait un GHJ de type PCTL et un de type PCL par sous-CMC).

Ces types de GHJ avaient l'inconvénient de manquer de signification sur la nature des pathologies prises en charge et présentaient pour certains des effectifs très petits. Le principe d'application de ce facteur a donc été modifié, en s'inspirant du modèle MCO, avec introduction d'un indicateur de surcoût lié aux actes médicaux ou aux diagnostics associés.

Ces coûts supplémentaires sont recherchés sur les 3 variables de morbidité principale (FP, MP, AE) mais aussi parmi les DAS, à condition qu'ils n'aient pas contribué au classement du recueil dans un GMD donné.

Parmi les évolutions envisageables de ce modèle, on pourrait envisager certaines segmentations pour homogénéiser les GMD (avec ou sans comorbidités), sachant que la recherche de la sévérité de la morbidité principale n'est pas la même chose que la mesure des effets des affections supplémentaires.

En ce qui concerne les actes médicaux, ils sont le plus souvent le signe d'un état sévère de la morbidité principale ; mais nous ne disposons pas de moyens automatiques permettant de déterminer le diagnostic ayant engendré la réalisation d'un acte donné.

Les listes de diagnostic et d'actes ont été établies à partir des données de l'ENC-SSR 2000 2001 en suivant la méthode suivante :

- pour les diagnostics, ont été retenus seuls les diagnostics associés non exclus (voir le paragraphe suivant),
- modélisation des coûts avec tous les autres facteurs, afin de déterminer le « résidu », à savoir la différence entre le coût mesuré et le coût estimé par le modèle, pour chaque RHA, la moyenne des résidus sur l'ensemble de la base étant par construction égale à zéro,
- détermination pour chaque acte CCAM, et chaque code diagnostic CIM10, de la moyenne des résidus des RHA possédant cet acte ou ce diagnostic ; cette moyenne ayant servi de mesure de « l'effet CAV ».

La détermination des CAV a été réalisée en combinant une expertise médicale (en particulier en tenant compte des PCL/PCTL) et les mesures de l'effet CAV définies précédemment, aboutissant à ne classer en CAV qu'un petit nombre de codes CIM10 et CCAM, et à ne faire qu'un seul niveau de sévérité (au lieu d'au moins deux comme pour les PCL/PCTL).

6.4.5 Le principe des exclusions

L'idée était de ne pas retenir comme comorbidités valorisantes (CAV) des codes diagnostics CIM10 qui présentent une certaine « proximité » avec les trois informations renseignées en morbidité principale (FP, MP, AE), ou même avec les diagnostics associés ayant servi à classer le recueil dans le GMD (en cas de double test). Comme pour le secteur MCO, on a considéré que la pondération supplémentaire apportée par une CAV devait résulter de la présence d'un facteur supplémentaire à celui décrit par le GMD.

Pour ce faire, ont été utilisés les derniers travaux MCO relatifs à l'élaboration de listes d'exclusion sur l'ensemble des diagnostics de la CIM10 (description des méthodes aboutissant à la constitution de ces listes d'exclusion dans le *rapport d'étape pour une version 11 des GHM*, publié en mai 2006⁽⁶⁾). L'utilisation pour le SSR d'un travail effectué en MCO se justifiait par le fait qu'une partie notable des exclusions correspondaient à du codage redondant et à des règles de codage CIM10 valables dans tous les secteurs. La liste des exclusions SSR relatives à un code CIM10 appartenant à la liste des CAV est détaillée dans l'annexe 4 du *manuel de groupage SSR 2009-1bis*⁽¹⁾.

Les exclusions assez nombreuses entre les pathologies voisines (du type « même sous paragraphe de la CIM10 ») peuvent pécher parfois par excès, notamment les couples de diagnostics en relation de cause à effet, par exemple les symptômes ou complications d'une pathologie pour lesquels il a été tenu compte le plus possible de l'impact supposé (à dire d'expert) et mesuré dans les bases. De plus, il n'est pas certain que les conclusions eussent été les mêmes en se plaçant dans le cadre SSR ; mais la validation en SSR nécessite une base de données de coûts journaliers assez conséquente, dont nous ne disposons pas à ce stade.

Toutes ces modifications ont permis d'élaborer une version IVA_{v0.2.}, utilisée dans le cadre des simulations réalisées au cours du 2^e semestre 2008, telle que :

$$\begin{aligned}
 & \text{IVA (PED=p, GMD=g, Sous-CMC=s, DepPhy=dp, DepCogn=c, NbReed5-12=r, Age=a, CAV=m,} \\
 & \quad \text{Finalp2=f, TypHosp=t)} \\
 & \quad = \\
 & \text{GMD(g) + DepPhy(g)*dp + DepCogn(s)*c + NbReed5-12(s)*r + Age(s,a) + CAV(s,m)} \\
 & \quad \text{+ Finalp2(s,f) + TypHosp(p,t)}
 \end{aligned}$$

Le nombre de points IVA journalier moyen pour un patient dans le GMD n^og, pour la sous CMCs, pour une séparation adultes/enfants=p, et ayant pour les autres facteurs les valeurs suivantes : Dépendance Physique=dp, Dépendance Cognitive=c, Nombre d'activité Cdarr =r, âge dans la classe n^oa, CAV=m, une finalité de prise en charge=f, et un type d'hospitalisation=t sera la somme de 8 termes.

Tableau D : Nombre de modalités possibles par variable composant l'IVA

Niveau de regroupement	Coefficient	Variations	Libellé
GMD	GMD(g)	Autant qu'il y a de GMD	
	DepPhy(g)	Autant qu'il y a de GMD	DepPhy(g) désigne le nombre de points de dépendance physique
Sous CMC	DepCogn(s)	Autant qu'il y a de Sous-CMC	DepCogn(s) désigne le nombre de points de dépendance cognitive
	NbReed(s)	Autant qu'il y a de Sous-CMC	NbReed(s) désigne le nombre de points par activité Cdarr
	Age(s,a)	Autant qu'il y a de Sous-CMC et de classes d'âge possibles	Désigne le nombre de points relatif à la classe d'âge a
	CAV(s,m)	Autant qu'il y a de Sous-CMC et suivant la présence d'une comorbidité ou d'un acte valorisant	Désigne le nombre de points relatif à la présence d'une CAV (si m=oui)
	Finalp2(s,f)	Autant qu'il y a de Sous-CMC et de modalités pour la variable Finalp	Désigne le nombre de points relatif à la Finalité principale=f
Distinction Adultes - enfants	TypHosp(p,t)	Quatre variations possibles	Désigne le coût relatif pour un type d'hospitalisation t soit chez un adulte, soit chez un enfant

La PVE du modèle v0.2. passe à 44.92% (pour 312 ddl).

6.5 Du modèle à son utilisation : la pondération en points IVA

Une fois l'étape de modélisation mathématique réalisée, le modèle de régression mis au point sur la base de l'optimisation de la PVE a permis de définir la pondération des données PMSI-SSR (RHA fournis par les établissements) sous la forme d'un nombre de points IVA (cf chapitre 3 du *Manuel de groupage SSR 2009-1bis⁽¹⁾*).

Un moteur informatique correspondant au modèle de groupage et de pondération, appelé « fonction groupage » a été développé (cf chapitre 4 du *Manuel de groupage SSR 2009-1bis⁽¹⁾*), à l'intention des logiciels de traitement des RHS (développés par l'ATIH) et des éditeurs pour intégration éventuelle dans les logiciels des établissements.

6.5.1 La formule

La formule relative à l'attribution de points IVA peut se présenter sous forme d'une matrice, chaque ligne correspondant à un GMD et chaque colonne à un facteur (publiée en annexe 7 du *manuel de groupage SSR 2009⁽¹⁾* et en annexe de ce rapport).

Prenons l'exemple d'un modèle relativement simple où le nombre de points IVA dépend du GMD, de la dépendance physique et de l'âge. La dépendance physique (score de 4-16 en principe) est utilisée comme variable quantitative. L'âge découpé en 3 classes : 18-44, 45-69 et 70 et plus (on prend des GMD « adulte »). La matrice des coefficients se présenterait alors sous la forme suivante :

Tableau E : Exemple simplifié de matrice de coefficients IVA

GMD	Base	Dep-Phys	âge		
			18-44	45-69	> 69
1	852	47	0	28	35
2	963	59	0	28	35
...					
35	1256	25	70	50	0
...					

Cet exemple n'utilise pas les valeurs réelles.

L'idée générale est que l'IVA pour une journée d'un recueil classé dans un GMD donné consiste à « additionner » les coefficients de la ligne correspondante, mais :

- en utilisant uniquement les coefficients adéquats (ceux qui correspondent aux caractéristiques du RHS). Sachant qu'il y a au minimum le coefficient de la colonne « base »,
- avec au préalable une multiplication quand la variable est utilisée en numérique. Par exemple, pour la dépendance physique, le nombre de points qu'elle induit est égal à la dépendance physique indiquée dans le RHA (à laquelle on soustrait la valeur 4 correspondant au niveau autonome), multipliée par le coefficient de la colonne « Dep-Phys ». Ce coefficient indique en fait le nombre de points IVA à rajouter par point de dépendance physique.

L'IVA pour une journée dans le GMD 1, avec une dépendance physique à 10 et d'âge de 50 ans sera pondérée à : $IVA = 852 + (10-4) * 47 + 28$.

Toujours dans le même GMD 1, un patient de 20 ans et de dépendance physique égale à 4 serait pondéré à 852 points.

6.5.2 Les classes de référence

Pour les variables mises en classes, telles que l'âge, une classe de référence (sans point ajouté) a été définie, pouvant varier d'un GMD à un autre.

Il serait donc possible de modifier les coefficients de la matrice, sans rien changer à la valeur finale de la formule. On peut toujours écrire que $972 = 972 + 25 - 25$; on peut rajouter une quantité X, par exemple à chaque coefficient des classes d'âge, à condition de soustraire X au coefficient de base (par exemple). Ainsi, la première ligne de la matrice donnée en exemple pourrait aussi s'écrire :

Tableau F : Exemple simplifié de matrice de coefficients IVA

GMD	Base	Dep-Phys	âge		
			18-44	45-69	> 69
1	752	47	100	128	135

Pour les variables en classes (pour un GMD donné), telles que « âge » ou « type d'hospitalisation », ce n'est donc pas tant la valeur absolue du coefficient qui veut dire quelque chose que sa différence avec les autres classes de la variable concernée. Par exemple, dans ces opérations d'addition/soustraction, la différence entre les deux classes d'âge 18-44 et 45-69 va toujours rester identique (28 points ici). Cette différence est appelée en statistique une « fonction estimable », qui ne change pas par convention d'écriture.

Il y a deux conventions d'écriture standard. La première consiste à déterminer les coefficients de manière à ce que leur moyenne (pour chaque variable en classe) soit nulle. Dans l'exemple du GMD 1, cela donnerait :

Tableau G : Exemple simplifié de matrice de coefficients IVA

GMD	Base	Dep-Phys	âge		
			18-44	45-69	> 69
1	873	47	-21	7	28

L'autre convention consiste à choisir une des modalités comme modalité de référence et à lui attribuer un coefficient nul. C'est ce qui a été retenu, en choisissant comme modalité de référence celle présentant le plus petit coefficient, de manière à ce que les autres soient positifs.

Pour les variables dichotomiques (ex : CAV sans/avec), le coefficient de la classe « sans » a été défini à zéro et seule la colonne relative à la pondération de la modalité « avec » figure dans la matrice.

6.5.3 Ajustement des coefficients

Une étape de validation des paramètres estimés (coefficients) a été ensuite réalisée, ayant conduit à supprimer ou regrouper un certain nombre de modalités lorsque les coefficients bruts étaient jugés non significatifs ou non pertinents.

Tableau H : Exemple de coefficients révisés des dépendances physiques

GMD	Effectif En RHA	Estimateur	Ecart-type	p-value	
12E05	22	15.7	31.9	0.62	Statistiquement non significatif, effectif en RHA très faible
13E21	803	0.0	7.3	0.99	Statistiquement non significatif
13E04	317	-13.7	9.6	0.15	Statistiquement non significatif
14E05	1003	-7.0	6.9	0.31	Statistiquement non significatif
15E01	136	-82.2	17.9	<0.001	Coefficient négatif
18E21	96	19.4	52.4	0.71	Statistiquement non significatif, effectif en RHA très faible
20E22	23	73.1	44.2	0.09	Statistiquement non significatif, effectif en RHA très faible, et tous les RHA avec dépendance élevée
50E21	88	-0.9	32.5	0.97	Statistiquement non significatif, effectif en RHA très faible

Pour ces situations, les coefficients ont donc été mis à zéro.

Pour les cas où tous les coefficients d'une même variable qualitative étaient d'emblée nuls, la variable a été considérée comme ne devant pas intervenir dans le GMD concerné pour deux raisons :

- la variable en question est jugée sans effet statistiquement significatif pour le GMD en question (ou la sous-CMC si le facteur n'a d'interaction qu'avec la sous-CMC),
- l'effet mesuré est peu compréhensible médicalement. Par exemple des effets de signe contraire à celui attendu a priori.

Lorsque tous les coefficients d'une même variable qualitative étaient égaux, cela revenait à neutraliser l'effet de la variable et à donc ne pas la retenir.

Enfin, si l'effet mesuré était peu compréhensible médicalement, il a semblé préférable de le neutraliser à cette étape. Pour exemple l'effet de l'âge en dent de scie : dans le modèle, seuls les cas présentant une évolution monotone (croissante ou décroissante), ou une courbe en U ou en U inversé ont été considérés comme acceptables. La plus grande partie de ces situations atypiques provenaient de sous-CMC avec effectifs faibles, pour lesquelles des regroupements de classe d'âge ont été réalisés.

Tableau I : Exemple de révision du coefficient « Age »

Niveau de regroupement		Age			
		19-29	30-59	60-74	75 et +
Sous CMC 40A	Estimations avant validation	374.0	107.7	114.7	0
	Effet en dents de scie				
	Regroupement de classes d'âge	19-29	30-74		75 et +
	Estimations après regroupement	373.6	90.4		0

Enfin, il a été nécessaire d'estimer les coefficients de certaines situations non représentées dans l'échantillon ENC (utilisé pour ces travaux), tels que :

- tous les coefficients relatifs au GMD 50E09, définis par analogie au GMD 50E08 et
- coefficients relatifs à la variable CAV des sous-CMC 16E et 18E à partir de ceux estimés pour les sous-CMC 16A et 18A.

6.5.4 Les interactions

La manière dont les effets d'un facteur (par exemple l'âge) varient en fonction d'autres facteurs (la sous-CMC ou le GMD par exemple) est déterminée par la notion d'interaction. On dit que deux facteurs X et Y sont sans interaction si l'effet de X est supposé ne pas dépendre de la valeur de Y. Si on impose dans le modèle que les effets de l'âge sont identiques pour tous les GMD d'une sous-CMC, cela se traduira par l'absence d'interactions entre l'âge et les GMD de la sous-CMC.

Dans le cas du GMD 17A03, par exemple (cf graphique C ci-dessous), on constate qu'il y a très peu d'interaction entre le score de dépendance physique et les classes d'âge.

Ce n'est pas parce que la ligne de chaque GMD contient une case pour « âge=45-69 » que toutes les valeurs dans cette colonne sont calculées de manière indépendante, et auraient donc toutes les chances d'être différentes (il y aurait eu autant d'estimation du coefficient « âge=45-69 » qu'il y a de GMD dans une sous-CMC donnée).

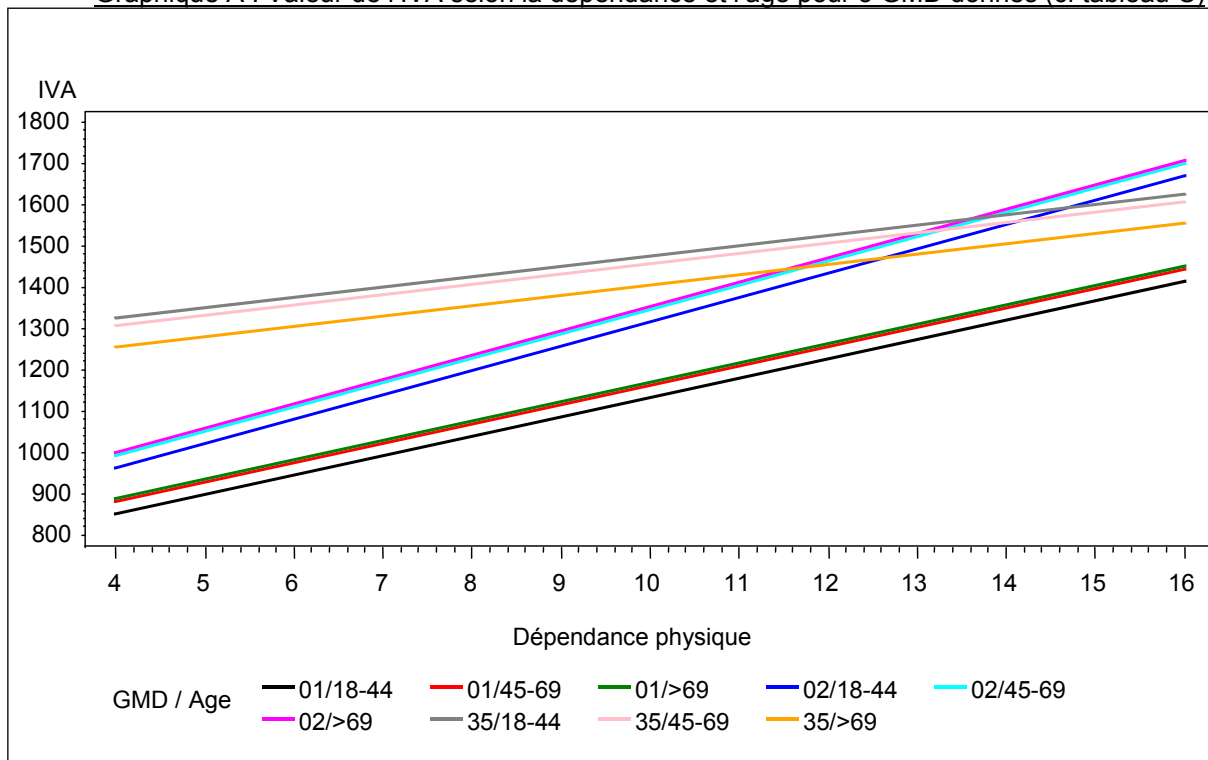
D'une manière générale, il n'était pas envisageable de déterminer avec une bonne robustesse les coefficients de cette matrice, compte-tenu du nombre de facteurs pris en compte, s'ils avaient été tous déclarés sans relation ; sans compter les effectifs insuffisants pour certains dans la base ENC-SSR.

C'est donc l'intérêt des modèles de régression que de pouvoir faire des hypothèses concernant la régularité des effets. Dans l'exemple donné, l'effet de l'âge est identique pour les deux premiers GMD ; ce qui n'est pas un hasard, le choix ayant été fait de contraindre l'effet « âge » par sous-CMC (soit calculer un coefficient identique pour tous les GMD d'une même sous-CMC).

Les hypothèses de travail ont été les suivantes :

- pas d'interaction à un niveau plus fin que le GMD. Par exemple, pour un GMD donné, l'effet de l'âge ne dépend pas du type d'hospitalisation ou du nombre de d'activité de rééducation-réadaptation, mais du GMD ;
- la dépendance physique étant le facteur principal expliquant les coûts, il y a donc eu intérêt à le prendre en compte de la manière la plus fine, c'est-à-dire :
 - o d'une manière continue, en supposant la relation coût x dépendance linéaire (ce qui a été vérifié)
 - o en faisant intervenir des interactions GMD x Dépendance, soit en estimant un coefficient spécifique de la dépendance pour chaque GMD ;
- pour les autres facteurs, le principe de parcimonie a été suivi et a consisté à ne pas retenir les interactions avec le GMD, pour se limiter à celles avec la sous-CMC (sauf lorsqu'il y avait une perte importante de PVE) ;
- pour les facteurs les plus rares, seules les interactions avec le facteur « pédiatrie/adulte » a été retenu, soit deux coefficients (un pour les enfants, un autre chez les adultes).

Graphique A : Valeur de l'IVA selon la dépendance et l'âge pour 3 GMD donnés (cf tableau C)



On constate que :

- l'effet de la dépendance est linéaire (ou plus exactement affine), à savoir représenté par une droite,
- que les 3 droites d'un GMD, fonction de l'âge, sont parallèles (mêmes pentes),
- que les pentes de chaque GMD, et leurs ordonnées à l'origine (coefficient de « base ») sont différentes.

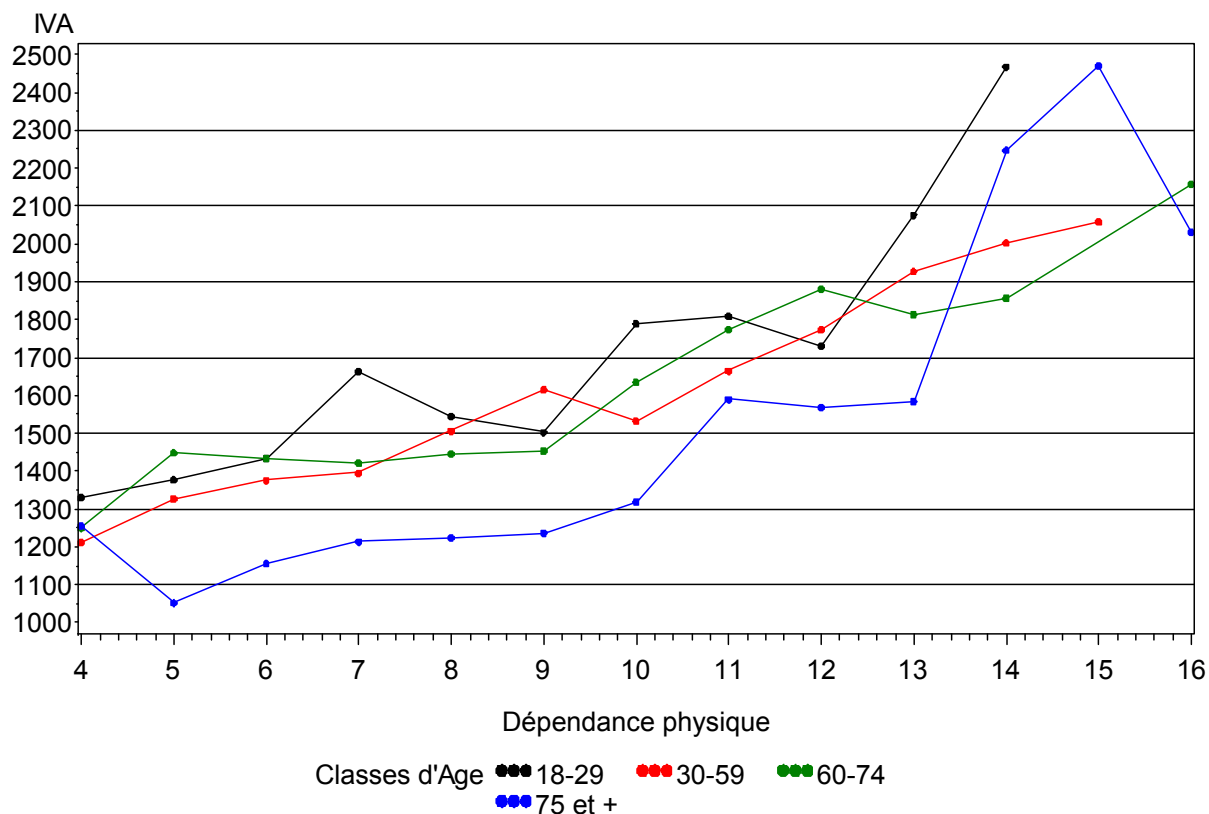
Par conséquent, un GMD peut être mieux valorisé qu'un autre pour certaines modalités de dépendances, et moins valorisé pour d'autres. Sur le graphique le GMD 02 part de plus bas que le GMD 35, mais présente des valeurs supérieures pour les dépendances élevées.

Dans le cas du GMD 17A03, par exemple, le graphique B ci-dessous illustre bien ce qui a été décrit de façon théorique dans le graphique A.

Graphique B

Nombre de points moyen par jour et par dépendance physique pour le GMD 17A03

GMD 17A03 : Lésions traumatiques multiples, complexes et compliquées, âge supérieur à 18 ans



Au final, la PVE obtenue est de 44.68% (pour 235 ddl).

7. Simulations et pertinence du modèle IVA

Pour poursuivre le modèle, il convenait de mettre au regard de cette production d'activité (en nombre de points IVA) une masse financière, dans le cadre d'une étape de simulation qui a été réalisée au cours du 2^e trimestre 2008 avec l'ensemble des ARH volontaires, tant sur le secteur public que privé, afin de s'assurer de la pertinence de l'IVA.

Pour ce faire, chaque tutelle régionale a été destinataire d'informations issues des données d'activité 2006 et a dû en retour se prononcer sur la pertinence des résultats propres à chacun de ses établissements SSR.

L'isolement des masses financières correspondant aux financements des différents compartiments du schéma cible a nécessité des travaux d'exploitation portant sur le retraitement comptable, sur l'enquête plateaux techniques spécialisés et sur les enquêtes relatives aux consommations de molécules onéreuses. En revanche, sur l'exercice 2006, ont été directement mobilisables les montants ayant financé l'activité SSR dans son intégralité. Ces montants ont donc pu être mis en regard des points IVA et ont ainsi permis de porter une première appréciation du modèle.

Cette simulation a porté sur un scénario utilisant les recettes complètes liées à l'activité connue, à partir de l'exercice 2006 ; la production en points IVA ayant été calculée à partir de la base nationale PMSI-SSR 2006 DGF et OQN.

Pour un exercice budgétaire donné, les recettes assurance maladie propres à l'activité de soins de suite et de réadaptation sont connues, telles que :

- pour les établissements « publics » cette activité est financée par dotation globale en dotation annuelle de financement. L'outil de suivi des campagnes budgétaires ARBUST permet de suivre proprement les dotations DAF. Au sein de la DAF, l'isolement des ressources finançant l'activité SSR est clairement spécifié pour les établissements soumis à la T2A. En revanche pour les établissements hors T2A, cet isolement a été rendu obligatoire à compter de 2007,
- pour les cliniques privées, la mobilisation des données de l'assurance maladie (SNIIRAM) permet de connaître pour chaque établissement les liquidations avec une possibilité de détail par code prestation et DMT.

Enfin il convient de noter que pour chacun des deux secteurs, l'appariement des données financières et des données PMSI s'est fait correctement à l'aide de l'établissement d'une liste de correspondance des FINESS.

Aussi, les données totales financières par établissement ont pu être mises en regard des données d'activité, en sachant que cette association ne tenait pas compte des compartiments spécifiques ni même de la spécificité du secteur OQN liée aux dépenses en sus.

Tableau J : Disponibilité des données PMSI-SSR 2006 pour simulation IVA

Exhaustivité source PMSI 2006 par Région	NB Etab DGF	% Etab / tot DGF	%Masse Financière / tot DGF	NB Etab OQN	% Etab / tot OQN	%Masse Financière / tot OQN
Alsace	39	84,80%	96,10%	4	100,00%	100,00%
Aquitaine	49	81,70%	93,80%	35	89,70%	97,00%
Auvergne	32	88,90%	96,50%	23	95,80%	99,00%
Bourgogne	39	92,90%	98,10%	18	94,70%	99,50%
Bretagne	73	98,60%	99,60%	6	100,00%	100,00%
Centre	39	76,50%	94,30%	18	94,70%	100,00%
Champagne-Ardennes	23	88,50%	94,90%	5	100,00%	100,00%
Corse	3	50,00%	49,60%	5	100,00%	100,00%
Franche-Comté	18	72,00%	86,80%	4	66,70%	98,00%
Ile-de-France	109	94,00%	98,50%	62	95,40%	97,00%
Languedoc-Roussillon	46	92,00%	96,80%	38	95,00%	100,00%
Limousin	18	90,00%	99,80%	2	100,00%	100,00%
Lorraine	58	90,60%	97,50%	8	88,90%	97,90%
Midi-Pyrénées	48	80,00%	87,90%	32	84,20%	96,70%
Nord-Pas-de-Calais	60	100,00%	100,00%	21	91,30%	97,10%
Normandie-Basse	25	64,10%	66,00%	12	85,70%	98,70%
Normandie-Haute	28	75,70%	93,40%	9	100,00%	100,00%
Pays-de-la-Loire	72	100,00%	100,00%	12	100,00%	100,00%
Picardie	41	93,20%	98,00%	4	100,00%	100,00%
Poitou-Charentes	31	86,10%	97,20%	9	100,00%	100,00%
PACA	77	97,50%	99,30%	85	96,60%	99,80%
Rhône-Alpes	129	94,90%	98,70%	23	95,80%	99,30%
Guadeloupe	4	57,10%	95,50%	2	22,20%	8,60%
Guyane	2	100,00%	100,00%	1	50,00%	83,10%
Martinique	7	77,80%	88,50%	2	66,70%	99,80%
Réunion	4	100,00%	100,00%	4	80,00%	100,00%

Au total, toutes les ARH ont souhaité participer à cette simulation, hormis la Guadeloupe (qui contenait 9 établissements OQN et 7 DGF en 2006) et la Guyane (1 OQN et 2 DGF), soit un taux de participation représentant 97% de la masse financière SSR DGF et 98% pour les OQN.

Tableau K : Périmètre de la simulation IVA

DG 2006	NB Etab	% Etab	%Masse Financière	OQN 2006	NB Etab	% Etab	%Masse Financière
Périmètre Simulation	1 072	89%	97%	Périmètre Simulation	444	92%	98%
Non transmission PMSI	129	11%	3%	Non transmission PMSI	39	8%	2%
TOTAL	1 201			TOTAL	483		

La méthode consistant à mettre en regard le nombre de points IVA produits et le financement 2006 de chaque établissement, une première étape s'est intéressée aux valeurs aberrantes des masses financières, soit 9 régions concernées de manière significative (> % national du secteur concerné) pour le secteur DGF et 11 pour le secteur.

Tableau L : Masses financières aberrantes

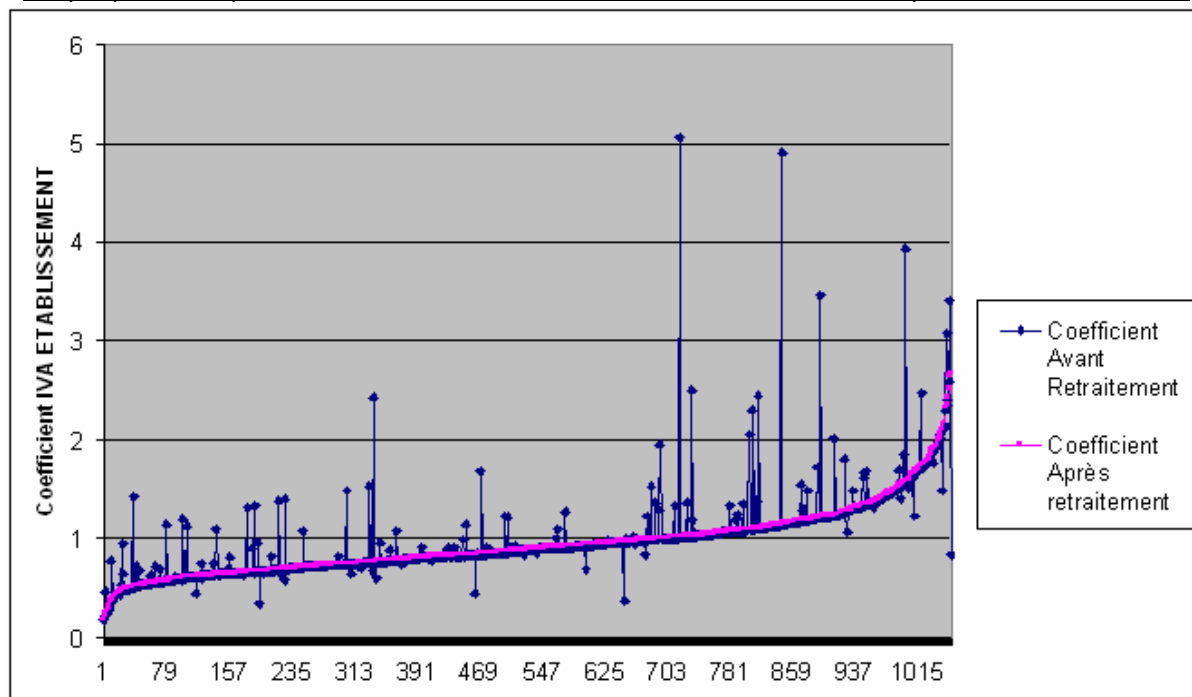
	DGF			OQN		
	NB Etab	% Etab	%Masse Financière	NB Etab	% Etab	%Masse Financière
Masse financière aberrante	116	10,86%	6,85%	102	23,02%	17,16%
dont étab T2A	30	25,9%	29,4%			

L'étape suivante consistait à signaler par les ARH toute incompatibilité entre le niveau d'activité transmis via le PMSI et la période couverte par le financement, soit 11 régions concernées de manière significative (> % national du secteur concerné) pour le secteur DGF et 6 pour le secteur OQN.

Tableau M : Incompatibilité activité et financement

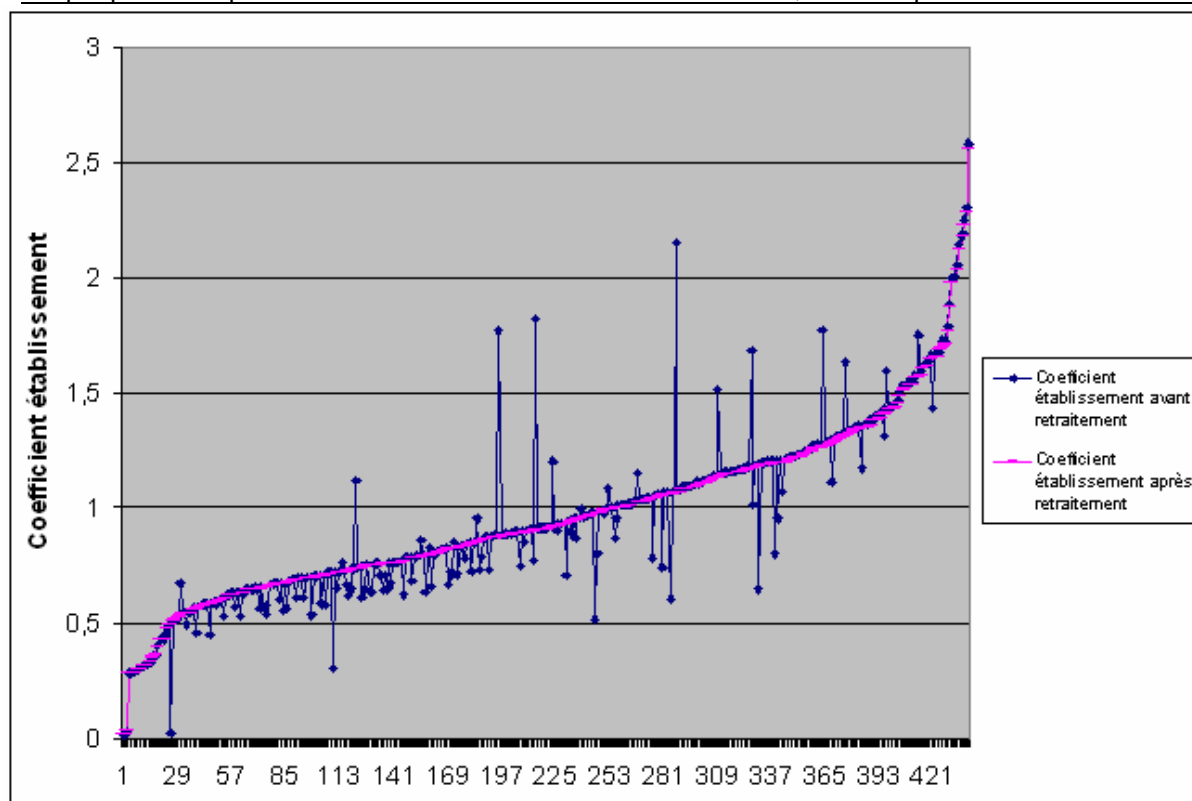
	DGF			OQN		
	Nombre étab	Pourcentage étab	Pourcentage masse financière	Nombre étab	Pourcentage étab	Pourcentage masse financière
Valeur point IVA recalculée suite modification (montant aberrant et/ou %PMSI) apportée par ARH	155	14,51%	9,97%	108	24,38%	18,19%
Incompatibilité activité/financement (modification %PMSI)	66	6,2%	5,4%	32	7,22%	2,15%

Graphique C : Répartition des coefficients IVA établissements DGF avant/après modifications ARH



$$\text{Coefficient IVA Etablissement} = \frac{\text{financement Etab/nombre de points IVA Etab}}{(\text{somme financement Etab/somme nbre points IVA Etab}) \text{ région concernée}}$$

Graphique D : Répartition des coefficients IVA établissements OQN avant/après modifications ARH



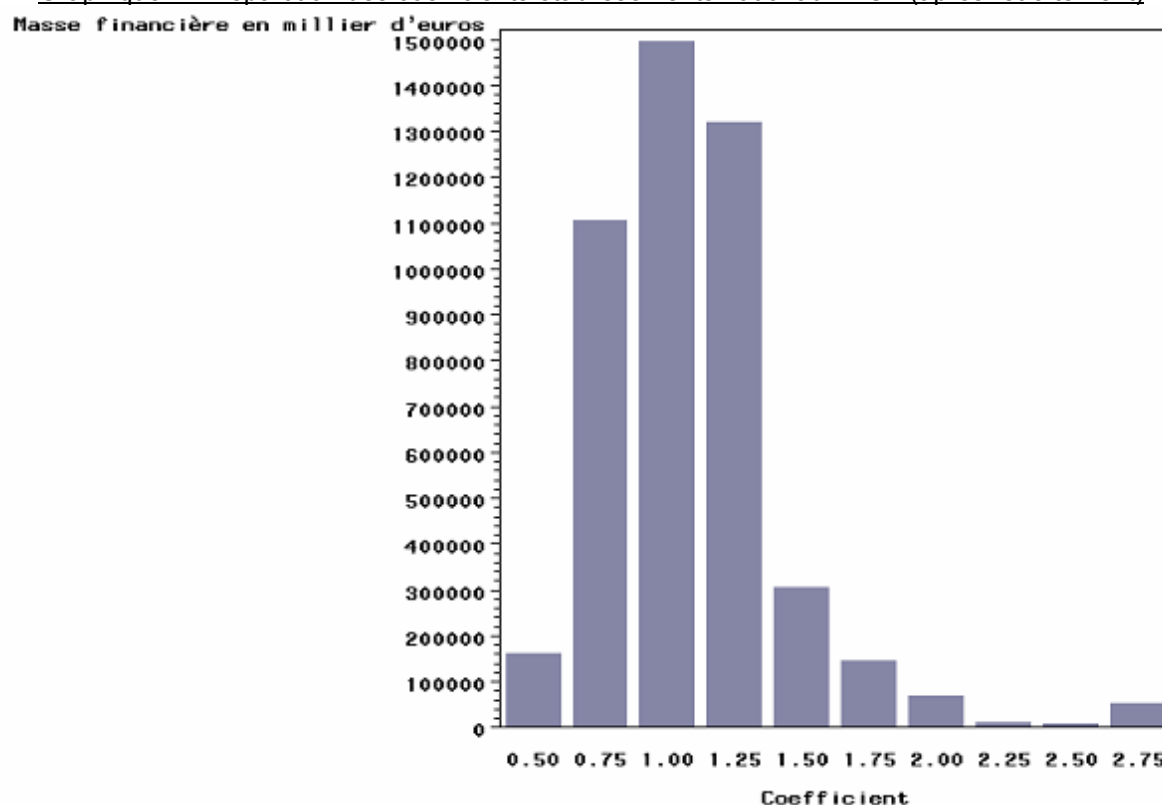
Une fois toutes les corrections nécessaires réalisées, il convenait pour les ARH de définir si l'application du modèle IVA reflétait correctement la position de chacun de leurs établissements (sur ou sous-doté).

Tableau N : Conclusions des ARH

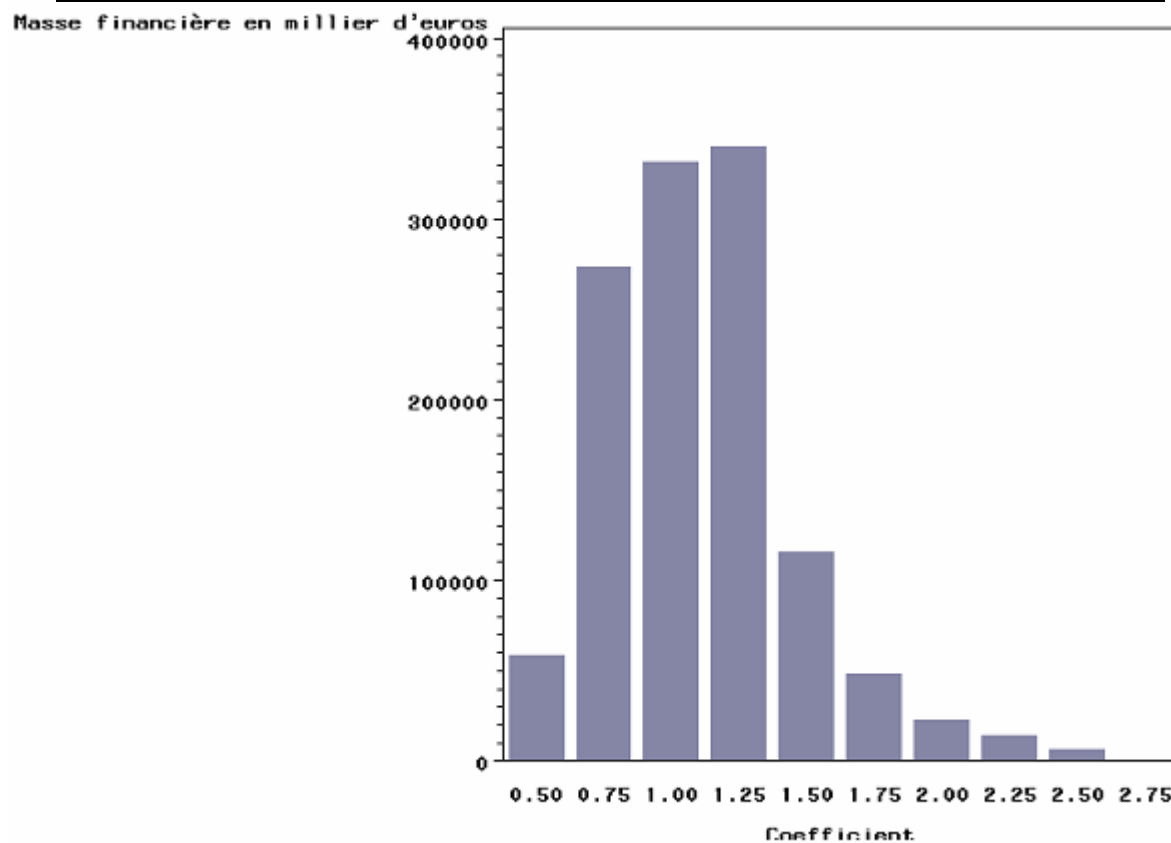
	DGF			OQN		
	Nombre étab	Pourcentage étab	Pourcentage masse financière	Nombre étab	Pourcentage étab	Pourcentage masse financière
Ne reflète pas position étab en 2006*	69	6,46%	4,84%	47	10,61%	6,31%
Non utilisation valeur point IVA 2006 (car situation particulière)	47	4,40%	4,77%	38	8,58%	4,43%
Evolution 2007-2008 susceptible de modifier position étab	148	13,86%	15,63%	58	13,09%	13,44%

Les réponses des ARH ont donc permis de conclure que l'IVA reflétait bien la position des établissements pour au moins 92% des structures et 95% de la masse financière.

Graphique E : Répartition des coefficients établissements nationaux DGF (après retraitement)

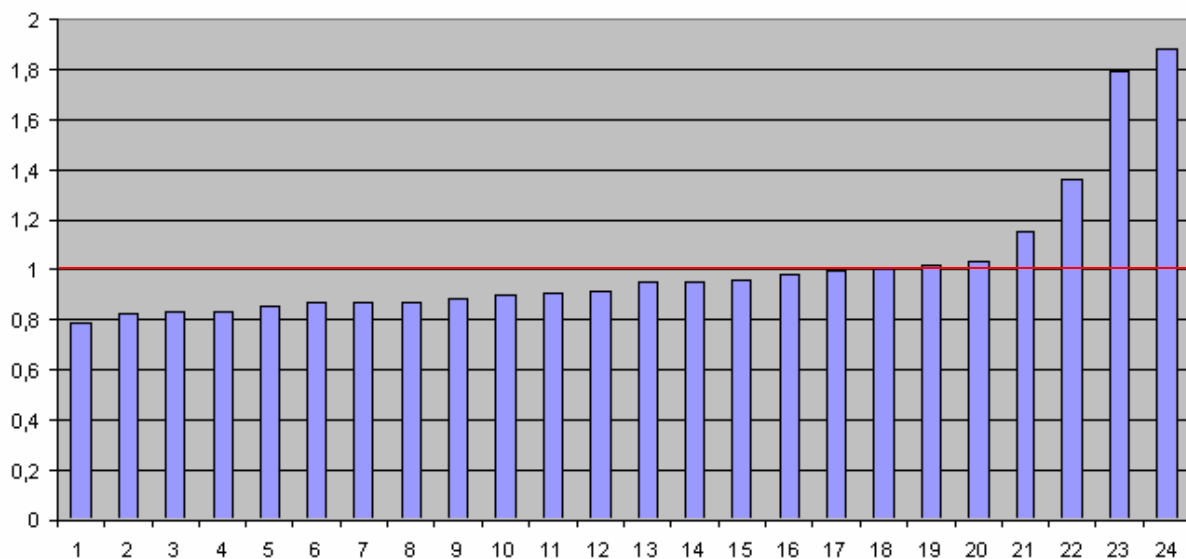


Graphique F : Répartition des coefficients établissements nationaux OQN (après retraitement)

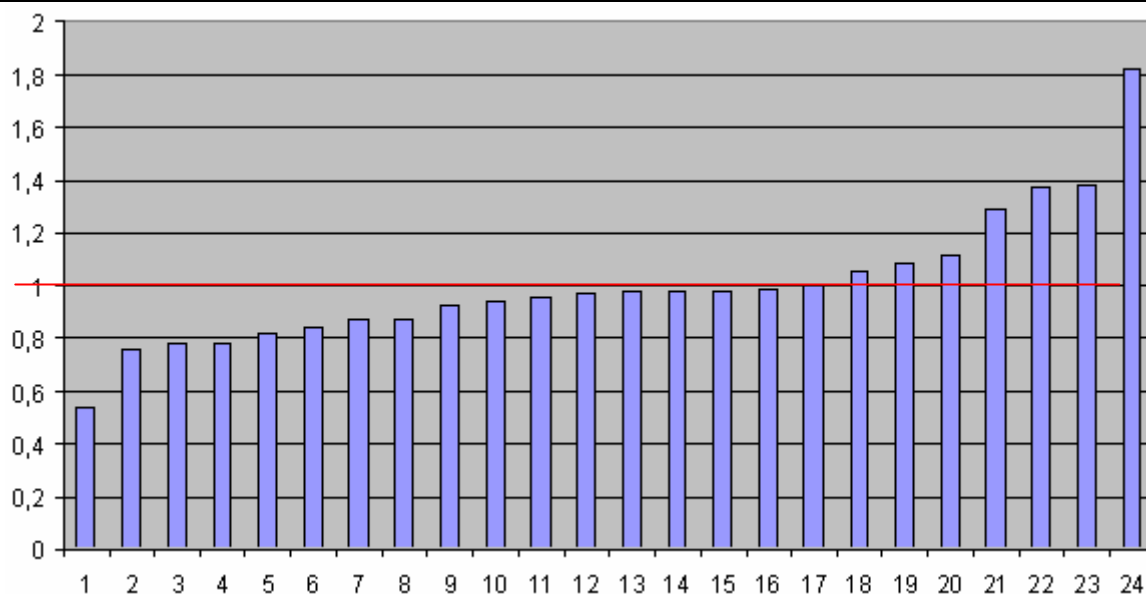


$$\text{Coefficient IVA Etablissement National} = \frac{\text{financement Etab} / \text{nombre de points IVA Etab}}{\left(\frac{\text{somme finance Etab} / \text{somme nbre points IVA Etab}}{\text{France entière}} \right)}$$

Graphique G : Moyenne régionale des coefficients établissement DGF / valeur nationale du point IVA



Graphique H : Moyenne régionale des coefficients établissement OQN / valeur nationale du point IVA



L'étude des moyennes régionales des coefficients établissement a fait apparaitre quatre régions concernant les établissements DGF et six pour les OQN avec des valeurs supérieures à 1,10 soit plutôt surdotées.

En conclusion, à ce stade, les régions n'ont pas signalé de difficultés majeures qui remettaient en cause l'utilisation de cet indicateur IVA et ont plutôt apporté une réponse favorable concernant sa pertinence en termes de positionnement des établissements.

Enfin, l'ATIH a également mené une étude de corrélation des variables composant l'IVA ainsi qu'une analyse factorielle, détaillées dans le document intitulé «Analyses IVA2007.pdf»⁷ (publié sur le site de l'ATIH), qui ont montré toutes deux que cet indicateur apparaissait pertinent en prenant bien en compte le poids de l'activité des établissements.

REFERENCES

-
- ¹ Manuel de groupage PMSI Soins de Suite et de Réadaptation, version 6 de la classification, version 4.6 de la fonction groupage - Bulletin Officiel N°2009-1bis fascicule spécial
- ² Manuel des groupes homogènes de journées, version 1 de la classification, version 1.1 de la fonction groupage, volumes I et II (Bulletin Officiel, Fascicule spécial n° 97/8 bis
- ³ Circulaire DHOS-PMSI-2001 n°105 du 22 février 2001 relative aux évolutions du PMSI pour l'activité de SSR des établissements financés par dotation globale
- ⁴ Étude nationale de coût PMSI SSR – exploitation du 2^e semestre 2000 relative aux établissements de santé sous dotation globale
Étude nationale de coût PMSI SSR – échelle 2^e édition – exploitation des données 2000/2001
- ⁵ Guide méthodologique de production des résumés hebdomadaires standardisés (Bulletin Officiel, Fascicule spécial n° 97/5 bis
- ⁶ Rapport d'étape pour une version 11 des GHM, mai 2006 (site ATIH – 2006)
- ⁷ Analyses IVA2007.pdf (site ATIH - 2009)