

Article original

Le programme de médicalisation du système d'information (PMSI) – processus de production des données, validité et sources d'erreurs dans le domaine de la morbidité maternelle sévère

French hospital discharge database: Data production, validity, and origins of errors in the field of severe maternal morbidity

A.A. Chantry^{a,*}, C. Deneux-Tharaux^{a,b}, G. Bal^c, J. Zeitlin^{a,b}, C. Quantin^{d,e},
M.-H. Bouvier-Colle^{a,b}, pour le groupe GRACE¹

^a *Inserm, UMR S953, recherche épidémiologique en santé périnatale, santé des femmes et des enfants, hôpital Saint-Vincent-de-Paul, 82, avenue Denfert-Rochereau, 75014 Paris, France*

^b *UMR S 953, UPMC, université Paris 06, 75005 Paris, France*

^c *Département de santé publique, CHRU de Grenoble, 38000 Grenoble, France*

^d *Inserm, U866, lipides, nutrition, cancer, 21000 Dijon, France*

^e *Université de Bourgogne, 21000 Dijon, France*

Reçu le 17 janvier 2011 ; accepté le 14 novembre 2011

Abstract

Background. – The organization of obstetric care in France brings all women in contact with the hospital system. Thus, hospital discharge data from the Program of Medicalization of the Information System (PMSI) constitute a potentially valuable source of information, particularly regarding rare events such as severe maternal morbidity. These data cover a large population but their quality has not been assessed in that field. Our objectives were to study the processes of production and the validity of PMSI data related to severe maternal morbidity.

Methods. – The study was conducted in four French tertiary teaching hospitals (Caen, Cochin [AP–HP, Paris], Grenoble and Lille). First, the organization of each step of the medical information process -production, formatting, verification and processing- was detailed in each center with a standardized form. Second, the validation study was based on the comparison of data related to severe maternal morbid events in the PMSI from these centers for 2006 and 2007, with the content of medical records which constituted the gold standard. Indicators of sensitivities and positive predictive values of PMSI were calculated.

Results. – The processes of PMSI data production showed major differences between the four centers. In hospital discharge data, diagnoses (eclampsia and pulmonary embolism) had a high proportion of false-positives (68%). Inversely, procedures (four procedures for management of severe haemorrhage) had less than 1% of false-positives, but a low sensitivity with 37% false-negatives which could be corrected in 95%. Regarding intensive care provision, all indicators of hospital data quality were very high. In addition, the validity of hospital data in centers 1 and 2 was higher for all events.

Conclusion. – The heterogeneity of the process of PMSI data production is associated with a variable quality of these data. Intensive care provision can be used in the PMSI, as well as procedures after correction. For diagnoses, the quality of the PMSI data is better in centers having both computerized medical records and steps for verification of medical information.

© 2012 Elsevier Masson SAS. All rights reserved.

Keywords: Acute severe maternal morbidity; Hospital discharge data; Validity; Sensitivity; Positive predictive value; Accuracy

* Auteur correspondant.

Adresse e-mail : anne.chantry@inserm.fr (A.A. Chantry).

¹ Groupe GRACE composé de : G. Bal, G. Beucher, M.-H. Bouvier-Colle, A.A. Chantry, C. Cans, M.-J. D'Alche Gautier, C. Deneux-Tharaux, A.-S. Ducloy-Bouthors, A. Ego, N. Lamendour, I. Le Fur, C. Quantin, V. Tessier, J. Zeitlin.

Résumé

Position du problème. – L'organisation des soins obstétricaux en France fait que les femmes rencontrent systématiquement le système hospitalier. Les données hospitalières nationales issues du programme de médicalisation du système d'information (PMSI) constituent donc potentiellement une source intéressante d'information, notamment pour l'étude d'événements rares tels que la morbidité maternelle sévère. Ces données couvrent un grand nombre d'individus mais leur qualité n'a pas été évaluée dans ce domaine. Notre objectif était d'étudier les circuits de production et la validité des données du PMSI relatives à la morbidité maternelle sévère.

Matériel et méthodes. – L'étude a été conduite dans quatre centres hospitaliers universitaires (Caen, Cochin, Grenoble, Lille). D'une part, l'organisation des étapes du circuit de l'information médicale – production, conditionnement, vérification et traitement – a été détaillée dans chaque centre grâce à l'utilisation d'une grille standardisée. D'autre part, l'étude de validation reposait sur la confrontation des événements maternels morbides sévères identifiés dans le PMSI des années 2006–2007 avec les données issues des dossiers médicaux qui constituaient la référence. Les sensibilités et valeurs prédictives positives du PMSI ont été calculées.

Résultats. – L'étude des circuits de production des données du PMSI révèle des différences majeures d'organisation entre centres surtout concernant l'étape de vérification de l'information médicale. Dans les données hospitalières, les diagnostics (éclampsie et embolie pulmonaire) présentaient une forte proportion de faux-positifs (68 %). À l'inverse, les actes (quatre actes marqueurs d'hémorragie sévère) avaient moins de 1 % de faux-positifs, mais une sensibilité basse avec 37 % de faux-négatifs dont 95 % étaient des erreurs corrigibles. Concernant les suppléments de soins intensifs et de réanimation, tous les indicateurs de validité étaient élevés. Par ailleurs, les centres 1 et 2 montraient des données de meilleure qualité quels que soient les événements.

Conclusion. – L'hétérogénéité des circuits de production des données du PMSI semble associée à une qualité variable de ces données. Les suppléments de réanimation et de soins intensifs, comme les actes après correction a posteriori peuvent être utilisés dans le PMSI. Pour les diagnostics, les données PMSI sont de meilleure qualité quand les centres possèdent conjointement une informatisation des dossiers et des étapes de vérification des informations médicales.

© 2012 Elsevier Masson SAS. Tous droits réservés.

Mots clés : Morbidité maternelle sévère ; PMSI ; Bases de données hospitalières ; Validation ; Sensibilité ; Valeur prédictive positive

1. Introduction

Pour évaluer leur activité, optimiser leur organisation et hiérarchiser leurs besoins, les établissements de santé se sont dotés d'un outil permettant de recueillir des données administratives et médicales [1–3]. Le programme de médicalisation du système d'information (PMSI), inspiré du modèle américain, collecte en routine et de façon pérenne toutes les données relatives aux séjours des patients dans les services hospitaliers [3,4]. Rapidement disponibles, peu coûteuses et riches en informations, les bases de données hospitalières comme le PMSI pourraient devenir des outils pertinents pour l'épidémiologie [4–6]. En obstétrique, il n'existe en France, actuellement, ni registre, ni cohorte permettant d'estimer la fréquence des complications de la grossesse, de l'accouchement ou de ses suites ; et la mortalité maternelle ne peut, du fait de sa rareté, demeurer le seul indicateur de santé maternelle des pays développés. Or, l'évolution des profils des mères (âge croissant à la maternité, augmentation de la prévalence de l'obésité et des maladies chroniques) et les modifications de pratiques obstétricales (augmentation des césariennes et des déclenchements) imposent d'étudier régulièrement les complications graves qui mettent en jeu le pronostic vital des femmes pendant la grossesse ou ses suites [7,8]. La fréquence de la morbidité maternelle sévère demeure pour autant complexe à estimer. Par exemple, depuis plus de dix ans, le groupe EURO-PERISTAT tente de recueillir des données sur ce sujet à partir de bases de données fiables et permanentes, sans y être encore parvenu [9].

Pourtant, chaque année en France, près de 830 000 femmes rencontrent systématiquement le système hospitalier pour l'accouchement ou lors de complications de la grossesse. Dans

ce contexte, les données hospitalières constituent potentiellement une source d'information utile, pérenne et une base pour détecter les événements rares de la morbidité maternelle.

Dans certains pays, elles sont employées depuis une vingtaine d'années pour évaluer les taux de complications suite à des hospitalisations [10], ou encore pour mesurer l'incidence de maladies [11,12]. Cependant, leur utilisation doit rester prudente. Les études de validation menées aux États-Unis [13,14], au Canada [10,15] ou en Australie [16–18] soulignent leurs nombreuses limites à savoir : une variabilité inter-établissement de la qualité du codage [19–21], des complications ou pathologies plus graves mieux codées [15,22], et des actes mieux saisis que les diagnostics [19,23]. En Europe, ces travaux sont peu nombreux et ont porté en France sur les domaines de la cancérologie [11,24,25], des soins intensifs [26] et du vasculaire [27], mais pas de l'obstétrique.

Les deux objectifs étaient d'étudier les circuits et la validité des données médicales obstétricales issues du PMSI relatives à la morbidité maternelle sévère.

2. Matériel et méthodes

2.1. Le programme de médicalisation des systèmes d'information (PMSI)

Le PMSI a été instauré en France en 1991 [1]. Après une phase d'expérimentation, il a été mis en place dans les établissements publics de santé dès 1994 [28], avant d'être généralisé en 1997 au secteur privé [29]. Conçu pour analyser l'activité des hôpitaux et contribuer à l'élaboration stratégique des projets d'établissements, le PMSI est devenu un instrument de gestion financière. Aux informations du PMSI ont été

attribués des tarifs et, dès 2004, la dotation globale des établissements publics a été progressivement remplacée par le système de tarification à l'activité (T2A) [30]. Dans le secteur privé, cette transformation s'est faite rapidement, puisqu'en un an (2005), ce nouveau mode de tarification couvrait 100 % de l'activité médicale des établissements de santé privés. Depuis 2008, le budget des établissements dépend entièrement de l'activité médicale décrite dans le PMSI [30]. Le PMSI compile les résumés standardisés de sortie (RSS : qui sont une synthèse des résumés d'unité médicale : RUM) de chaque séjour enregistré dans l'établissement de santé. Secondairement, les RSS sont anonymisés et deviennent les résumés de sortie anonymisés (RSA). Les informations qu'ils contiennent sont d'ordre administratif (âge, sexe, code géographique de résidence, année, mois et mode d'entrée, année, mois et mode de sortie du séjour, statut de l'établissement) et d'ordre médical [28]. Les diagnostics reconnus lors du séjour sont codés selon la Classification internationale des maladies dans sa dixième révision (CIM10 depuis 1993). Tous les actes médico-techniques effectués pendant le séjour sont codés selon la Classification commune des actes médicaux (CCAM). Les règles du PMSI sont nationales et dictées par les tutelles : ministère de la santé et Agence de traitement de l'information hospitalière (ATIH) [31]. Chaque établissement produit ses données anonymes standardisées, lesquelles sont compilées au niveau régional, puis national. Notre étude porte sur la validation de certains événements maternels morbides sévères dans la base nationale du PMSI. Dans cette base, les informations sont présentées sous forme de RSA. Chaque ligne décrit un séjour et se définit par une suite de codes qui correspondent en colonne aux variables administratives et médicales successivement présentées. Pour décrire les actes par exemple, plusieurs colonnes sont utiles : outre celle comportant le code de l'acte en sept caractères, une autre décrit le nombre de réalisation de l'acte ou encore le délai de réalisation de l'acte depuis la date d'entrée du séjour.

2.2. Types d'étude

Il s'agit d'une étude descriptive comportant deux investigations menées en parallèle. La première est une observation de l'organisation des systèmes d'information hospitaliers. La seconde est une validation des données issues du PMSI relatives à la morbidité maternelle sévère en référence aux informations issues des dossiers médicaux des patientes.

2.3. Population d'étude

Le travail a été mené dans quatre centres hospitaliers universitaires français (Caen, Cochin [AP-HP, Paris], Grenoble et Lille), choisis pour leurs maternités de type III qui concentrent les cas les plus sévères de morbidité liée à la grossesse ou ses suites. La participation des établissements s'est faite sur la base du volontariat et selon trois critères :

- équilibre Île-de-France/Province ;

- volonté des équipes cliniques et celles des départements d'information médicale d'initier un travail de recherche sur la qualité de leur PMSI ;
- utilisation par l'établissement d'un dossier médical informatisé.

Nous avons sélectionné dans la base nationale du PMSI les séjours des femmes en âge de procréer (14–50 ans inclus), ayant accouché dans les maternités étudiées et dont la date de sortie était comprise entre le 1^{er} janvier 2006 et le 31 décembre 2007. L'accouchement était identifié par au moins un code de la CIM 10 parmi O80.0 à O84.9 et Z37.0 à Z37.9. Les femmes n'ayant pas accouché dans les maternités étudiées ont été exclues car leurs dossiers médicaux n'étaient pas complets.

2.4. Collecte des données et stratégie d'analyse

Deux catégories de données ont été recueillies.

D'une part, pour mener les observations relatives à l'organisation des systèmes d'information hospitaliers (SIH) des quatre centres, une personne du groupe (A.C.) s'est déplacée dans chacun d'eux. Une grille standardisée a permis de recueillir les informations concernant chaque étape du circuit de l'information médicale : de sa production à son traitement. Les informations relevées concernaient les types de support de l'information médicale, l'informatisation des données, le personnel dédié à l'étape de vérification de l'information médicale, la revue des dossiers et les normes de standardisation utilisées. Les quatre étapes retenues pour l'observation étaient les suivantes :

- première étape : production primaire de l'information médicale source. À ce stade, l'information médicale reconnue par les soignants est transcrite dans le dossier médical du patient et représente l'information médicale source. La production primaire de l'information médicale par les soignants marque son introduction dans le SIH ;
- deuxième étape : conditionnement de l'information médicale et production du PMSI. Ici, l'information médicale source inscrite dans le dossier médical du patient subit des transformations. Elle est conditionnée pour s'intégrer dans une base de données structurée. Les règles de codage sont des normes nationales, mais l'interprétation des informations demeure de la responsabilité du codeur ;
- troisième étape : vérification et validation de l'information médicale. L'objectif de cette étape est d'assurer la conformité de l'information médicale produite avec les données médicales sources. En cas de détection d'erreurs dans l'information médicale produite, c'est à ce stade que sont apportées les corrections ;
- quatrième étape : traitement de l'information médicale finale. Avant transmission des données du PMSI aux tutelles (ARS, ATIH), le personnel spécialisé (techniciens d'information médicale et médecins de santé publique) des départements d'information médicale des établissements effectue une ultime vérification de l'information médicale en repérant et traitant les incohérences et les atypies du PMSI.

D'autre part, l'étude de validation, dont la méthode est détaillée ailleurs [32], a porté sur les événements maternels morbides sévères (EMMS) les plus fréquemment cause de décès maternels et qui signent par nature la sévérité [33]. Une approche des EMMS à travers diverses catégories de codage (diagnostic, acte, élément de prise en charge) a été privilégiée afin d'étudier la validité de chacune de ces catégories. Il s'agit de : deux diagnostics : éclampsie (code CIM10 : O15) et embolie pulmonaire (code CIM10 : O88) ; quatre actes pratiqués dans le cadre d'hémorragies du post-partum : embolisation (code CCAM : EDSF011), ligature des artères utérines (code CCAM : EDSA002), ligature des pédicules vasculaires utérins (code CCAM : ELSA002) et hystérectomie d'hémostase (code CCAM : JNFA001) ; et deux suppléments de prise en charge : supplément de soins intensifs (code PMSI : SUP STF) et de réanimation (code PMSI : SUP REA). Les séjours comportant au moins un des EMMS considérés étaient sélectionnés dans le PMSI grâce à la recherche des codes spécifiques des EMMS, et dans les dossiers médicaux grâce à leurs libellés. Quand plusieurs résumés de sortie décrivaient le même EMMS pour la même femme, l'EMMS était compté une seule fois.

Les données relatives aux EMMS issues du PMSI ont été confrontées aux données du dossier médical des patientes. La sélection des EMMS dans les dossiers médicaux a été rendue possible par l'interrogation d'une base de données complémentaire : la base de données des dossiers médicaux informatisés. Ces fichiers étaient disponibles dans les quatre centres. Les EMMS ont été identifiés dans les bases de dossiers informatisés par recherche de leurs libellés et des variantes de leurs libellés (écritures différentes, acronymes, noms des EMMS au singulier et au pluriel). Nous avons fait le choix de cette méthode car il était long et onéreux de lire les informations des plus de 30 000 dossiers médicaux de la population d'étude. Le dossier médical était considéré comme la référence.

Les dossiers médicaux informatisés étaient reliés aux séjours du PMSI par les variables suivantes disponibles dans les deux bases de données : âge de la femme, mois et mode d'entrée, mois et mode de sortie de séjour, durée du séjour et code géographique de résidence de la patiente.

La confrontation consistait en une lecture du dossier médical source et de toutes ses composantes de synthèse finale : lettres de sortie, dossiers infirmiers et comptes-rendus hospitalier et opératoire. Aucune interprétation d'examens, ni aucun jugement sur les diagnostics n'ont été portés a posteriori. L'EMMS était inscrit dans le dossier médical ou ne l'était pas.

Dans ce contexte, les vrais positifs représentaient ainsi les EMMS identifiés à la fois dans les résumés du PMSI et dans les dossiers médicaux correspondants. À l'inverse, les faux-positifs étaient les événements enregistrés dans le PMSI qui n'existaient pas comme tels dans les dossiers des patientes. Les faux-négatifs représentaient les EMMS inscrits dans les dossiers médicaux des patientes, mais non saisis dans le PMSI. Au contraire, les vrais négatifs représentaient les situations pour lesquelles à l'absence d'EMMS dans les dossiers médicaux

sources des patientes correspondait bien le non-enregistrement de l'EMMS dans le PMSI.

Dans les centres 1 et 3, le dossier médical informatisé était confondu avec le dossier médical source ; le dossier papier n'étant qu'une impression de son contenu. Dans les centres 2 et 4, le dossier médical source était le dossier papier, et les informations étaient secondairement saisies dans le dossier informatisé, quotidiennement, par les cliniciens, au fil de l'hospitalisation des patientes.

L'accord de la Commission nationale de l'informatique et des libertés avait été obtenu (n° 1004749).

2.5. Analyses

Une description des circuits d'information des quatre centres est proposée.

Pour estimer la qualité de l'enregistrement des EMMS dans le PMSI, la sensibilité, et la valeur prédictive positive (VPP) du PMSI en référence au dossier médical source ont été estimées à l'aide du logiciel Stata[®] version 10. La sensibilité indique la probabilité du PMSI d'enregistrer correctement une femme présentant un EMMS. La VPP correspond à la probabilité qu'une femme ait présenté effectivement un EMMS quand cet événement est identifié dans le PMSI.

3. Résultats

3.1. Description des circuits d'information

Les observations menées dans les quatre établissements retenus ont montré de nombreuses disparités dans l'organisation de leur SIH (Tableau 1).

Lors de l'étape de production primaire de l'information médicale, dans tous les centres, un large éventail de professionnels, aux compétences diverses, était impliqué dans la production de l'information médicale source : infirmières, externes, internes, sages-femmes, chefs de clinique, médecins et professeurs.

L'information médicale existait dans les établissements, sous des formats hétérogènes. Alors que dans les centres 1 et 3, le dossier médical source se présentait exclusivement sous forme informatisée (le dossier papier n'est qu'une impression du dossier informatisé), les centres 2 et 4 possédaient un dossier médical papier doublé d'un format informatique. Chacun des établissements utilisait son propre logiciel de gestion informatique de l'information médicale.

Concernant l'étape de conditionnement de l'information médicale, les méthodes utilisées étaient hétérogènes, malgré des règles de codage nationales. Dans les centres 1 et 3, les praticiens renseignaient les informations médicales dans les dossiers informatisés au fur et à mesure du séjour des patientes et ces informations étaient immédiatement et automatiquement conditionnées et intégrées dans le PMSI, selon une programmation pré-établie. Dans le centre 4, à l'issue du séjour de la patiente, les praticiens saisissaient sous format informatique toute l'information médicale sous forme de codes CCAM et CIM10 à partir du dossier médical papier, en se référant

Tableau 1
Description des étapes composant le système d'information hospitalier par centre.

	Centre 1	Centre 2	Centre 3	Centre 4
<i>n</i> accouchements sur la période 2006–2007	6555	10 486	3970	9596
<i>Étape 1 : production de l'information primaire</i>				
Support de l'information médicale				
Dossier papier				
Dossier informatisé	X		X	
Les deux formats		X		X
<i>Étape 2 : conditionnement de l'information</i>				
Règles nationales				
<i>Étape 3 : vérification de l'information standardisée</i>				
Étape de vérification				
Oui	X	X		X
Non			X	
Si vérification, nombre de personnes pour cette tâche				
1–2	X	X		
≥ 3				X
Si vérification, qualification du personnel dédié à cette tâche				
Infirmières				X
Sages-femmes	X	X		X
Si vérification des dossiers				
Tous les dossiers	X			
Revue partielle		X		X
Si revue partielle				
Les dossiers complexes uniquement (pathologies, transferts. . .)		X		
Une partie du séjour uniquement				X
<i>Étape 4 : traitement de l'information médicale finale</i>				
Département d'information médicale (DIM)				
Centralisé	X	X	X	
Décentralisé				X
Équivalent temps plein (ETP) du personnel du DIM				
ETP de médecin	2,8	2	3,5	1 ^a
ETP de technicien d'information médicale	4	4	8	1 ^a
Démarche qualité				
Contrôle de l'exhaustivité des séjours	X	X	X	X ^a
Vérification des séjours atypiques (durées trop longues, erreurs, incohérences)	X	X	X	X ^a
Vérification de procédures particulières (ex. césarienne)		X		

^a Ces observations concernent le DIM du pôle enfant et gynécologie-obstétrique seulement.

notamment à des thesaurus proposés dans le service. Le cas du centre 2 était différent puisqu'il combinait les deux méthodes. Une partie du codage de l'information médicale s'effectuait au fur et à mesure du séjour de la patiente, directement grâce aux informations contenues dans le dossier informatisé, tandis qu'à la sortie, les praticiens codaient les informations complémentaires manquantes.

Au sein des quatre établissements, l'étape présentant le plus d'hétérogénéité était celle de la vérification de l'information médicale. Dans les centres 1, 2 et 4, l'information médicale subissait une vérification avant son envoi au département d'information médicale (DIM). Des sages-femmes s'attachaient à cette fonction, avec dans chaque centre, des activités distinctes. Le centre 1 disposait d'une unique sage-femme à temps plein (un équivalent temps plein : ETP), qui procédait à une revue complète de l'information médicale de tous les séjours obstétricaux de l'établissement, incluant les séjours de post-partum. À l'inverse, dans le centre 2, l'unique sage-femme dédiée à cette activité (0,5 ETP), n'effectuait qu'une revue partielle des séjours les plus à risque d'erreurs (transferts

d'autres établissements, accouchements à domicile, cas de grossesses multiples ou hospitalisations répétées). Le centre 4, employait à temps partiel quatre titulaires à cet effet (2,5 ETP) qui n'effectuaient qu'une revue partielle de l'information médicale. Ces personnes vérifiaient uniquement les données des séjours d'hospitalisations de grossesse à risque et d'accouchements. Les informations concernant le post-partum n'étaient pas vérifiées. Le centre 3 n'avait pas intégré cette étape dans son SIH, considérant que la responsabilité de la validité de l'information médicale produite était celle de chaque intervenant.

Les procédures de traitement de l'information médicale finale étaient identiques dans les quatre DIM. Tous avaient initié une démarche qualité concernant les données issues du PMSI. Celle-ci consistait en un contrôle de l'exhaustivité des séjours et en une vérification de certains séjours : ceux aux durées trop longues, ceux groupés en erreur, et ceux présentant des atypies de codage ou des incohérences. Le centre 2 avait instauré une vérification supplémentaire de la cohérence entre certains actes et diagnostics d'un même séjour, pour la

césarienne par exemple. L'organisation des DIM et leur allocation en personnel différaient selon les centres. Cela est détaillé dans le **Tableau 1**. Les centres 1, 2 et 3 possédaient un DIM centralisé pour tout l'hôpital et disposaient d'au moins six ETP (médecins et techniciens confondus). À l'inverse, le DIM du centre 4 avait un fonctionnement décentralisé, avec une antenne implantée dans chaque pôle de l'hôpital. Dans celle du pôle enfant et gynécologie-obstétrique, l'équipe comprenait un ETP de médecin et un ETP de technicien d'information médicale.

3.2. Étude de validation des données issues du PMSI en référence à celles du dossier médical

Parmi les 30 614 femmes ayant accouché dans les quatre centres pour les années 2006 et 2007, 403 EMMS ont été identifiés dans le PMSI. Sept cas ont été exclus de l'analyse, car les dossiers médicaux sources correspondants n'ont pas été retrouvés. Ainsi, l'analyse a porté sur 396 EMMS identifiés dans le PMSI, correspondant à 318 femmes sur une base de 30 607 accouchements.

L'analyse des faux-positifs et des faux-négatifs dans le PMSI a conduit à distinguer trois situations : pour les diagnostics une proportion élevée de faux-positifs ; au contraire, concernant les actes, une proportion élevée de faux-négatifs ; et pour les suppléments, peu de faux-positifs et faux-négatifs (**Tableau 2**). Avec un taux de 67,8 %, les diagnostics présentaient une proportion élevée de faux-positifs dans le PMSI. Soixante-dix-huit des 115 diagnostics repérés dans le PMSI n'étaient pas validés par la lecture des dossiers médicaux. C'est par exemple le cas de 67 éclampsies sur 84 repérées dans le PMSI qui étaient en réalité des pré-éclampsies sévères ou des protéinuries isolées (tableaux cliniques différents et globalement de gravité moindre).

Sur 129 actes pratiqués dans le cadre d'hémorragies du post-partum, à l'exception de l'un d'entre eux, il n'existait pas de faux-positifs. En revanche, la proportion de faux-négatifs atteignait 36,9 %. Ce sont, au total, 75 actes qui n'étaient pas enregistrés correctement dans le PMSI. Les actes avaient été pratiqués et codés sans prise en compte du contexte post-partum, alors qu'un code précis existe pour ce contexte spécifique. La CCAM propose par exemple pour l'hystérectomie une vingtaine de

codes distincts. Ainsi l'hystérectomie totale par laparotomie est codée JKFA015, alors que le même geste réalisé dans le cadre spécifique d'une hémorragie du post-partum, a pour code approprié JNFA001.

Les suppléments de réanimation et de soins intensifs comptabilisaient autant d'évènements identifiés dans les données issues du PMSI que dans les dossiers médicaux, malgré 2 % de faux-positifs et 2 % de faux-négatifs repérés dans les deux fichiers.

Les analyses par centre montraient des résultats contrastés (**Tableau 2**). Les EMMS des centres 1 et 2 semblaient avoir été codés correctement dans le PMSI avec des faibles proportions de faux-positifs respectivement de 2,7 % et 5,6 % et de faux-négatifs de 2,7 % et 15,6 %. Inversement, dans les deux autres centres étudiés, des erreurs de codage avaient été mises en évidence. Le centre 3 indiquait un taux de faux-négatifs atteignant 57,6 %, et le centre 4 présentait un taux important de faux-positifs de 43,0 %.

Le **Tableau 3** présente les sensibilités et VPP du PMSI en référence au dossier médical par type d'EMMS et par centre. Ces indicateurs atteignaient des valeurs élevées pour les suppléments de réanimation et de soins intensifs, supérieurs à 96 % dans tous les centres. La catégorie des diagnostics présentait pour un centre, une VPP du PMSI élevée (83 %) ; pour les trois autres centres, la VPP variait de 13 à 69 %. En revanche, les sensibilités des diagnostics dans le PMSI, étaient élevées, elles étaient supérieures à 83 % dans trois centres sur quatre. À l'inverse, les actes étaient caractérisés par des VPP du PMSI par centre très élevées, proches de 100 %, tandis que leurs sensibilités, variant de 18 à 94 %, reflétaient les nombreux faux-négatifs retrouvés dans les dossiers médicaux des patientes. Ces faux-négatifs correspondaient à des actes codés dans le PMSI sous des codes généraux, au lieu de l'être sous les codes spécifiques de la grossesse, l'accouchement ou de ses suites. Nous avons considéré ces erreurs comme rectifiables quand le reste des éléments du séjour de la patiente indiquait un contexte gravidopuerpéral. Après correction sous l'hypothèse que tous ces actes étaient réalisés dans le contexte de la grossesse considérée, les analyses révisées aboutissaient à des sensibilités supérieures à 92 % dans tous les centres.

La **Fig. 1a** et **b** résument graphiquement les relations entre informatisation du dossier médical, implication de

Tableau 2

Évènements maternels morbides sévères (EMMS) identifiés dans le PMSI et les dossiers médicaux des patientes et pourcentages de faux-positifs (FP) et faux-négatifs (FN) dans le PMSI^a.

	n dans PMSI	n dans dossiers médicaux	% FP	% FN
<i>n évènements totaux</i>	396	399		
Diagnostics	115	44	67,8	15,9
Actes	129	203	0,8	36,9
Suppléments de réanimation/soins intensifs	152	152	2,0	2,0
<i>n évènements totaux</i>	396	399		
Centre 1	74	74	2,7	2,7
Centre 2	126	141	5,6	15,6
Centre 3	38	66	10,5	48,5
Centre 4	158	118	43,0	23,7

^a Sur une base de 30 607 accouchements, en prenant le dossier médical des patientes comme référence.

Tableau 3

Validité des données issues du PMSI relatives aux évènements maternels morbides (EMMS) : sensibilité et valeur prédictive positive (VPP)^a.

		Centre 1	Centre 2	Centre 3	Centre 4
Diagnostics	Sensibilité (%)	100,0	73,3	85,7	83,3
	VPP (%)	83,3	68,7	60,0	13,0
Actes	Sensibilité (%)	93,9	77,9	18,4	13,0
	Sensibilité révisée ^b (%)	93,9	97,4	97,4	54,5
	VPP (%)	100,0	98,4	100,0	100,0
	VPP révisée ^b (%)	100,0	98,7	100,0	100,0
Suppléments de réanimation/soins intensifs	Sensibilité (%)	100,0	97,9	100,0	96,1
	VPP (%)	100,0	97,9	100,0	100,0

^a Sur une base de 30 607 accouchements, en prenant le dossier médical des patientes comme référence.^b Résultats révisés après correction des codes d'actes non spécifiques au contexte obstétrical.

l'établissement dans l'étape de vérification de l'information médicale à travers les critères de revue des dossiers et du nombre de personnes impliquées, et validation des données du PMSI. Le centre 3, par exemple, présentait la proportion la plus importante de faux-négatifs, et était le seul établissement qui ne comprenait pas d'étape intermédiaire de vérification de l'information médicale obstétricale.

Le centre 4 en revanche, comptabilisait le plus de faux-positifs. Ce centre disposait de 2,5 ETP de sages-femmes dédiées à la vérification des données médicales qui effectuaient une revue partielle des dossiers (dossiers d'hospitalisations de grossesses à risque et d'accouchements mais pas de post-partum). En revanche, il ne disposait pas d'un SIH très développé et le dossier médical n'était que partiellement informatisé et peu utilisé au quotidien par les cliniciens. À l'inverse, les centres 1 et 2, avec une informatisation développée des dossiers, une revue plus importante des dossiers, mais moins de personnel dédié à cette tâche (1 et 0,5 ETP respectivement), présentaient les données PMSI les plus valides de l'étude. Ces observations semblaient par ailleurs indépendantes de la taille des maternités, basée sur leur nombre annuel d'accouchements.

4. Discussion

L'étude des circuits d'information médicale montre des différences majeures entre les quatre établissements, tant en termes d'organisation de la production d'information qu'en termes de vérification de la qualité de l'information produite. Les choix d'investissement dans ces circuits sont très hétérogènes selon les établissements. Il est aussi intéressant d'observer que la qualité des données issues du PMSI relatives à la morbidité maternelle sévère est contrastée selon les catégories d'informations (diagnostics, actes et suppléments) et variable entre les centres.

La question de la validation des données hospitalières est fondamentale pour leur utilisation en épidémiologie. Les cliniciens doutent de la qualité des données issues du PMSI et le faible nombre de travaux qui s'intéressent à leur validation ne les encourage pas à les utiliser à des fins de recherche. Quelques équipes françaises ont entamé des travaux de validation du PMSI dans certains domaines mais ces études sont longues, coûteuses et complexes, notamment concernant la question

centrale des données de référence (registre, dossier du patient, dossier recodé par des experts ou étude prospective).

Une particularité de notre étude est d'avoir cherché à décrire le lien entre l'organisation des circuits de production de l'information médicale et la qualité de cette information. L'association qui semble exister entre, d'une part, l'investissement simultané d'un établissement dans l'informatisation du dossier médical et dans l'étape intermédiaire de vérification des informations et, d'autre part, la qualité des données du PMSI, permet de proposer des pistes aux établissements désireux d'améliorer la qualité de leur PMSI. À ce jour, nous ne connaissons pas de travaux publiés sur cet aspect.

Les résultats de notre étude suggèrent que l'utilisation du PMSI est prometteuse pour obtenir des informations essentielles en obstétrique, sur les actes pratiqués (à quelques corrections près) et les suppléments de réanimation et de soins intensifs, en cas d'EMMS.

Ces suppléments ne présentent pas uniquement des informations administratives caractérisant l'admission dans une unité médicale, mais attestent indirectement de la gravité des états cliniques des patientes et de leur prise en charge. À titre d'exemple, le supplément réanimation en 2006/2007 correspondait à l'admission d'un patient en service de réanimation ayant un indice de gravité simplifié supérieur ou égal à 15 et pour lequel trois actes de la liste II ou un acte de la liste I de l'arrêté du 5 mars 2006 ont été pratiqués. Par ailleurs les logiciels de codage produisent ces suppléments grâce à une mécanique de génération automatique dès qu'ils détectent les éléments qui les définissent. Le codage dans le PMSI de ces suppléments de soins intensifs et de réanimation est moins soumis à la subjectivité du codeur ; leurs très faibles proportions de faux-positifs et faux-négatifs nous permettent de considérer leur codage dans le PMSI comme précis et fiable. Ayant montré leur validité dans le contexte obstétrical, nous pouvons envisager leur utilisation à plus large échelle.

Toutefois notre étude comporte certaines limites.

Tout d'abord, elle porte sur la validité du PMSI dans seulement quatre établissements tous hospitalo-universitaires. Même si les complications de la période gravidopuerpérale sont parfois traitées ailleurs que dans les hôpitaux universitaires, ces centres concentrent les EMMS au niveau régional. Ils représentent donc des établissements adaptés pour une

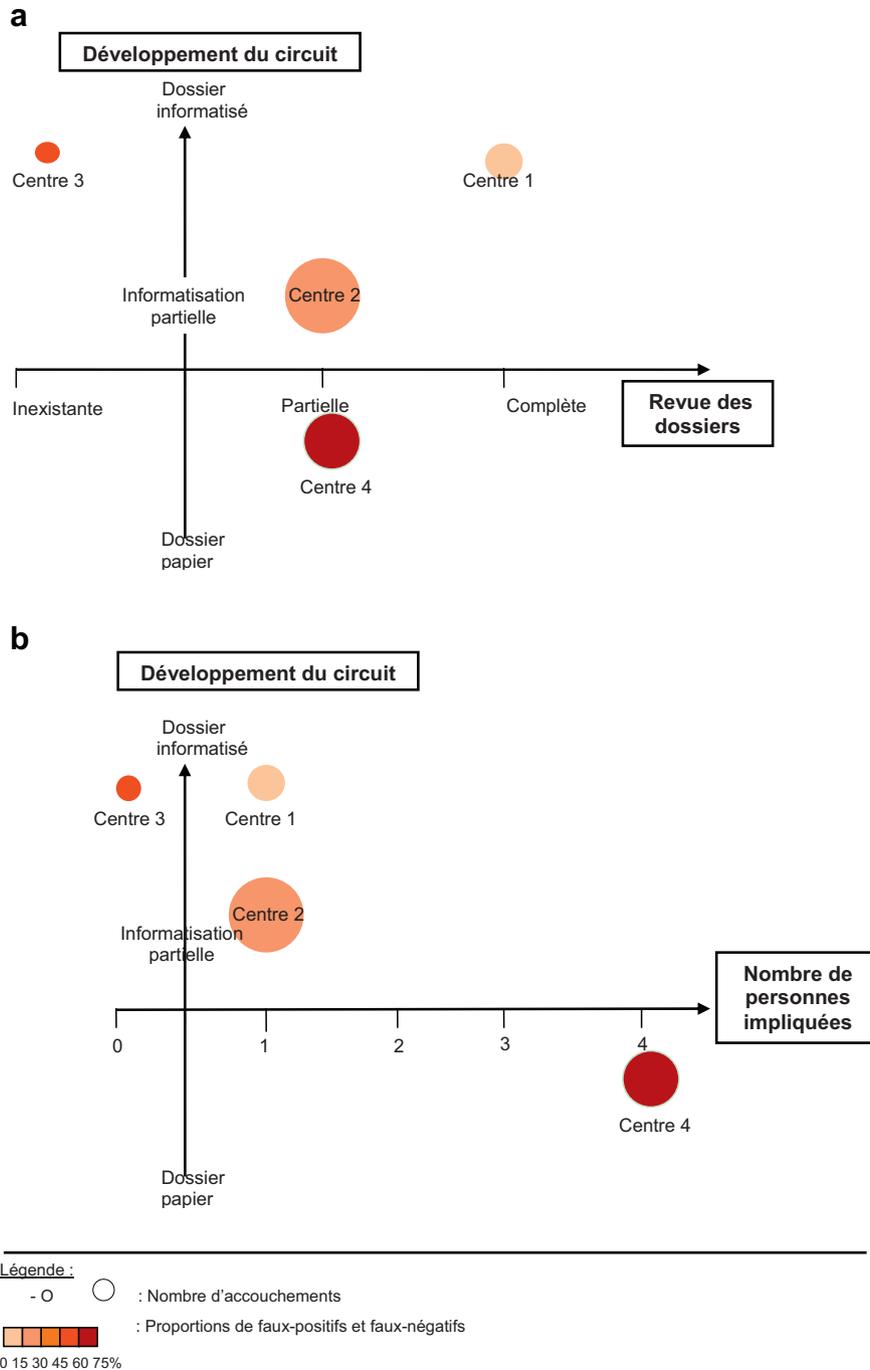


Fig. 1. Typologie des circuits d'information médicale : a : revue des dossiers lors de l'étape de vérification de l'information médicale ; b : nombre de personnes investies dans l'étape de vérification de l'information médicale.

première étude du PMSI dans le domaine de la morbidité maternelle sévère. Il n'est pourtant pas exclu que quelques EMMS soient pris en charge dans d'autres catégories d'établissements. Une étude similaire menée dans des établissements privés ou non universitaires dont l'organisation des SIH présenterait des spécificités serait intéressante pour améliorer la connaissance. Nous envisageons de la mener prochainement.

Le nombre restreint d'établissements étudiés ne permet pas la généralisation des résultats concernant certains EMMS à

l'échelon national, jusqu'à plus ample information. Cependant, cette limite est de moindre portée concernant les suppléments de réanimation et de soins intensifs puisqu'ils sont très majoritairement implantés dans les hôpitaux universitaires ou régionaux.

Les EMMS sélectionnés ne couvrent pas toute la morbidité maternelle sévère. La définition de celle-ci n'étant pas consensuelle sur le plan international, nous avons retenu les EMMS les plus fréquemment cause de décès maternels et qui signent par nature la sévérité.

Enfin, la méthode de validation employée peut être discutée. Notre objectif était de valider les données issues du PMSI relatives aux EMMS. Plus précisément, nous avons cherché à savoir si le PMSI retranscrivait bien les actes pratiqués et les diagnostics retenus par l'équipe en charge de la patiente. Ainsi, à la différence des auteurs qui ont étudié la validité des données hospitalières avec un gold standard basé sur une réinterprétation en insu des données du dossier médical [10,13,20,22,23,26,34–40], notre référence reposait sur les données contenues dans les dossiers, c'est-à-dire, sur les diagnostics retenus par l'équipe et qui ont justifié la prise en charge de la patiente. Sans réinterprétation, ils reflétaient la réalité des SIH au quotidien.

L'étude combinée des circuits de l'information médicale et de la validité des données PMSI issues de ces circuits nous a permis de relever certains facteurs impliqués dans les mécanismes menant à des erreurs dans les bases de données hospitalières. Nous discutons ici, le caractère réversible des divers types d'erreurs et l'implication des équipes dans les étapes de correction ; et tentons d'identifier secondairement les causes induisant dans le PMSI des faux-positifs et faux-négatifs.

4.1. Réversibilité des erreurs

Certaines erreurs sont définitives et ne peuvent être corrigées. C'est le cas de celles issues de la non saisie d'un évènement dans le dossier médical source, ou encore de certaines erreurs d'ordre informatique. A contrario, les erreurs de codage ou dues à l'absence de codage dans la base informatisée ou dans le résumé de séjour peuvent être corrigées, d'où l'intérêt de les connaître.

C'est, dans les centres étudiés, tout le travail effectué par le personnel dédié aux étapes de vérification des données, et des personnels des DIM plus tard dans le circuit. Sur ce point, le groupement pour la modernisation du système d'information hospitalier insiste sur le facteur ressources humaines qui demeure un point majeur de la qualité des systèmes d'information hospitaliers [41]. Les erreurs de codage, telles que les imprécisions ou surqualifications, peuvent être réversibles dans le circuit de l'information, à condition que du personnel qualifié soit attaché à cette fonction précise. L'informatisation des circuits de production des données hospitalières et le stade de vérification de l'information médicale semblent être des éléments clés pour élaborer un PMSI homogène, exhaustif et fiable [42]. Cependant, nos observations suggèrent, non seulement l'importance de la mise en place d'une étape de vérification des informations avec un nombre adéquat de personnes allouées à cette fonction, mais aussi la définition de leurs tâches. Ainsi, la complétude de la revue des dossiers serait un facteur favorisant une information de qualité, comme le montrent les résultats du centre 1.

Certaines erreurs à défaut d'être réversibles dans la base de données finale, peuvent être corrigées a posteriori au moment des analyses. C'est le cas des erreurs dues à des codages imprécis liés au contexte du post-partum non spécifié. Quand pour ces actes, le reste des éléments du séjour du patient indique un contexte gravido-puerpéral, des corrections peuvent être appliquées. Les calculs révisés aboutissent alors à des sensibilités très élevées

(> 92 %), pour ces actes, dans tous les centres. Dans ces conditions, il nous paraît tout à fait acceptable d'étudier la fréquence de certains actes dans le PMSI.

4.2. Les faux-positifs

Les faux-positifs identifiés dans le PMSI correspondent aux évènements repérés dans les bases de données hospitalières pour lesquels le dossier médical ne donne pas la même information.

Deux mécanismes principaux d'erreurs de codage induisent des faux-positifs dans le PMSI en référence au dossier médical. Tout d'abord, la source d'erreur de codage peut être l'utilisation d'un autre code. Erreur de sélection dans un menu déroulant, erreur de lecture, ou d'inattention, dans notre étude, moins de 2 % des faux-positifs sont dus à l'utilisation d'un code pour décrire une information qui n'est pas le bon. L'autre source d'erreur de codage induisant des faux-positifs peut être l'utilisation d'un code par excès. Dans ce cas, le codage traduit une situation médicale plus sévère que celle décrite dans le dossier médical [42,43]. Nos résultats montrent qu'un codage en excès a été identifié pour près de 70 % des diagnostics.

L'étude approfondie des cas sur-codés d'éclampsies rapportait qu'il s'agissait en réalité de pathologies hypertensives globalement moins graves ; de même pour les embolies pulmonaires, les faux-positifs étaient majoritairement dans les dossiers médicaux des thromboses veineuses profondes. Ces erreurs de codage ne constituent pas une exception. D'autres auteurs retrouvaient des VPP pour l'éclampsie dans les bases de données hospitalières variant de 23,5 % dans une étude multicentrique australienne [18,34] à 41,7 % pour une étude monocentrique à Chicago [44], ou encore de 50 % pour une étude dans l'état de Washington.

Dans la littérature, plusieurs arguments expliquent les causes possibles de surqualification des diagnostics dans les bases de données hospitalières :

- le large panel de codeurs intervenant dans le codage qui entraîne une hétérogénéité de qualité de l'information médicale [21,35,36,42,45,46] ;
- les cas les plus graves ayant entraîné la mobilisation de toute l'équipe médicale sont surestimés au moment du codage pour indiquer le caractère sévère du tableau clinique [34,36,47,48] ;
- le système de tarification basé sur la gravité des séjours est un incitatif fort à des excès de codage [4,6,13,45].

Ainsi, certaines pathologies peuvent être surqualifiées et aggravées en comparaison aux pathologies observées pour valoriser le séjour du patient et permettre à l'hôpital de percevoir un remboursement plus important. C'est un biais du système de tarification décrit depuis longtemps [4,6,13,45,49].

4.3. Les faux-négatifs

Les faux-négatifs dans le PMSI correspondent aux évènements vécus par les patientes et décrits dans leurs

dossiers médicaux mais non codés dans les bases de données hospitalières. Ils concernent généralement les actes.

Dans le processus de production de l'information hospitalière, trois mécanismes peuvent induire des faux-négatifs. Le premier est l'absence de codage. L'information médicale peut ne pas avoir été reportée dans le dossier médical de la patiente ; elle est dans ce cas, non produite et n'existe pas dans le SIH. Ou alors, les informations sont décrites dans les dossiers médicaux des patientes, mais non saisies dans la base informatisée des dossiers médicaux et non codées dans le PMSI. Dans notre étude, ce problème apparaît dans 5 % des cas. Deuxièmement, le mécanisme induisant des faux-négatifs dans le PMSI peut être l'erreur de codage. Ces erreurs de codage peuvent être de deux ordres : l'utilisation d'un mauvais code ou de codes imprécis. Aussi, dans ce cas d'imprécision de codage, l'acte est codé, mais le code utilisé ne traduit pas le contexte spécifique dans lequel il a été réalisé [42,43]. Troisièmement, les erreurs peuvent être dues plus rarement à des problèmes techniques. La défaillance des systèmes informatiques entraînant alors des pertes d'informations à tous les niveaux.

Dans notre étude, 95 % des erreurs induisant des faux-négatifs étaient des erreurs d'imprécision de codes. On retrouvait, ainsi, à la lecture des dossiers médicaux concernés, des actes réalisés en urgence pour pallier une hémorragie du post-partum codés dans le PMSI comme des actes programmés hors contexte obstétrical. Les codes particuliers reliés à des contextes spécifiques n'étaient pas pris en compte. Une étude multicentrique australienne dressait le même constat en retrouvant pour les hystérectomies réalisées dans le cadre d'hémorragies de la délivrance, une sensibilité de 28,3 %, attribuant ce résultat aux erreurs de codes spécifiques [18,22].

Les causes des imprécisions de codage induisant des faux-négatifs évoquées dès les années 1990 [14,42,45,48] et encore aujourd'hui [22,23] sont la conséquence de l'utilisation de classifications lourdes à manipuler. Le codage devient complexe, la CCAM propose des codes d'actes de plus en plus précis avec des codes distincts pour décrire un même acte réalisé dans des contextes différents. Pour les cliniciens, le codage devient de plus en plus chronophage et ils enregistrent alors dans les bases hospitalières, les actes avec le code qu'ils utilisent le plus souvent sans rechercher de façon systématique les codes les plus appropriés. L'utilisation de thesaurus est aussi une cause d'erreur induisant des faux-négatifs. Fiches résumées des codes d'actes réalisés le plus régulièrement dans les services, les thesaurus facilitent le codage mais ne décrivent pas les situations rares et sévères [23,42].

En pratique, dans le domaine des complications obstétricales, si le codage des suppléments de soins intensifs et de réanimation dans le PMSI est valide, celui des actes et des diagnostics peut être encore amélioré par des actions simples.

La première consisterait à organiser une étape de vérification systématique des données avant leur production finale. Si la revue complète de tous les dossiers par une seule personne paraît complexe à organiser, une revue ciblée de dossiers particuliers peut être envisagée. La qualité des données ne semble pas dépendante du nombre de personnes attachées à la vérification des informations, mais l'homogénéité de leurs

compétences paraît importante. De même, la fenêtre temporelle de la revue des dossiers est un élément important. Les données sont de meilleure qualité quand la revue de l'information médicale couvre toute la période gravido-puerpérale et n'exclue pas la période du post-partum.

De même, l'informatisation complète du système d'information hospitalier améliorerait la fiabilité du système. C'est tout le travail des informaticiens qui, par ce biais, pourraient développer des outils informatiques de correction automatique à partir d'algorithmes, ou plus simplement, introduire dans les systèmes informatiques des mécanismes de feed-back avec des fenêtres de contrôle du codage, avant validation d'une information.

On peut également recommander une amélioration reposant sur le traitement statistique de l'information. Ainsi, la surveillance d'indicateurs tels que les actes réalisés, par exemple pour traiter l'hémorragie sévère du post-partum, s'améliorerait par l'élargissement des codes utilisés, quand ils décrivent un séjour gravido-puerpéral.

Quoi qu'il en soit, l'amélioration future du système est prévisible. Le remboursement des prestations étant directement corrélé aux informations du PMSI, l'assurance maladie effectue des contrôles pour relever les abus et soumettre les établissements au remboursement des prestations injustifiées. La multiplication de ces contrôles externes, additionnée aux efforts des personnels des DIM qui effectuent régulièrement des contrôles internes devraient entraîner une amélioration générale de la qualité du PMSI.

5. Conclusion

Certains facteurs simultanément appliqués favoriseraient la qualité des données PMSI, tels que : l'utilisation d'un dossier médical informatisé et la vérification par du personnel dédié du conditionnement des données du PMSI dans leur intégralité à l'issue des séjours.

Le codage, dans le PMSI, des suppléments de réanimation et de soins intensifs est de bonne qualité. Les actes effectués dans le cadre d'hémorragie du post-partum peuvent être utilisés pour l'épidémiologie quand les codes spécifiques au contexte et codes généraux sont combinés ; tout comme les diagnostics peuvent être utilisés dans les centres où une étape de vérification de l'information médicale est bien organisée et couvre toute la période gravido-puerpérale. Dans ces conditions, l'utilisation à l'avenir des bases de données médico-administratives ouvre des perspectives pour la surveillance de la morbidité maternelle sévère.

Déclaration d'intérêts

Les auteurs déclarent ne pas avoir de conflits d'intérêts en relation avec cet article.

Remerciements

Nous remercions mesdames Annick Blondel, Maité Meriaux et Julie Tort pour leur participation.

Cette recherche a bénéficié de l'aide conjointe du ministère de la Santé et des Solidarités, Direction générale de la santé, de l'Institut national de la santé et de la recherche médicale, de l'Agence française de sécurité sanitaire des produits de santé et de l'Institut de veille sanitaire, dans le cadre de l'appel à projets de recherche lancé par l'IRESP en 2006.

Références

- [1] Ministère de la Santé. Loi n° 94-748 du 31 juillet 1991 portant réforme hospitalière. In: Journal Officiel de la République française, n° 179 du 2 août 1991. Paris; 1991: 10255.
- [2] Ministère de la Santé. Décret n° 94-666 du 27 juillet 1994 relatif aux systèmes d'information médicale et à l'analyse de l'activité des établissements de santé publics et privés. In: Journal Officiel de la République française, n° 180 du 5 août 1994. Paris; 1994: 11395.
- [3] Fetter RB, Shin Y, Freeman JL, Averill RF, Thompson JD. Case mix definition by diagnosis-related groups. *Med Care* 1980;18(2 Suppl.):iii. 1–53.
- [4] Lombrail P, Minvielle E, Comar L, Gottot S. Programme de Medicalisation des Systèmes d'Information et épidémiologie : une liaison qui ne va pas de soi. *Rev Epidemiol Sante Publique* 1994;42(4):334–44.
- [5] Kone Pefoyo AJ, Rivard M, Laurier C. Public health surveillance and role of administrative data. *Rev Epidemiol Sante Publique* 2009;57(2): 99–111.
- [6] Geoffroy-Perez B, Imbernon E, Gilg Soit Ilg A, Goldberg M. Confrontation des données du Programme national de surveillance du mésothéliome (PNSM) et du Programme de médicalisation du système d'information (PMSI). *Rev Epidemiol Sante Publique* 2006;54(6):475–83.
- [7] Brace V, Penney G, Hall M. Quantifying severe maternal morbidity: a Scottish population study. *BJOG* 2004;111(5):481–4.
- [8] Say L, Souza JP, Pattinson RC. Maternal near miss—towards a standard tool for monitoring quality of maternal health care. *Best Pract Res Clin Obstet Gynaecol* 2009;23(3):287–96.
- [9] EURO-PERISTAT Report Writing Committee: European Perinatal Health Report. 2008:280.
- [10] Quan H, Parsons GA, Ghali WA. Assessing accuracy of diagnosis-type indicators for flagging complications in administrative data. *J Clin Epidemiol* 2004;57(4):366–72.
- [11] Couris CM, Polazzi S, Olive F, Remontet L, Bossard N, Gomez F, et al. Breast cancer incidence using administrative data: correction with sensitivity and specificity. *J Clin Epidemiol* 2009;62(6):660–6.
- [12] Mahonen M, Salomaa V, Keskimaki I, Moltchanov V, Torppa J, Molarius A, et al. The feasibility of combining data from routine Hospital Discharge and Causes-of-Death Registers for epidemiological studies on stroke. *Eur J Epidemiol* 2000;16(9):815–7.
- [13] Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. *N Engl J Med* 1988;318(6):352–5.
- [14] Lloyd SS, Rissing JP. Physician and coding errors in patient records. *JAMA* 1985;254(10):1330–6.
- [15] Quan H, Parsons GA, Ghali WA. Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. *Med Care* 2004;42(8):801–9.
- [16] Taylor LK, Travis S, Pym M, Olive E, Henderson-Smart DJ. How useful are hospital morbidity data for monitoring conditions occurring in the perinatal period? *Aust N Z J Obstet Gynaecol* 2005;45(1):36–41.
- [17] Hadfield RM, Lain SJ, Cameron CA, Bell JC, Morris JM, Roberts CL. The prevalence of maternal medical conditions during pregnancy and a validation of their reporting in hospital discharge data. *Aust N Z J Obstet Gynaecol* 2008;48(1):78–82.
- [18] Roberts CL, Cameron CA, Bell JC, Algert CS, Morris JM. Measuring maternal morbidity in routinely collected health data: development and validation of a maternal morbidity outcome indicator. *Med Care* 2008; 46(8):786–94.
- [19] Fisher ES, Whaley FS, Krushat WM, Malenka DJ, Fleming C, Baron JA, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health* 1992;82(2):243–8.
- [20] Iezzoni LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E, Levine PA. Coding of acute myocardial infarction. Clinical and policy implications. *Ann Intern Med* 1988;109(9):745–51.
- [21] Smulian JC, Ananth CV, Hanley ML, Knuppel RA, Donlen J, Kruse L. New Jersey's electronic birth certificate program: variations in data sources. *Am J Public Health* 2001;91(5):814–6.
- [22] Lain SJ, Roberts CL, Hadfield RM, Bell JC, Morris JM. How accurate is the reporting of obstetric haemorrhage in hospital discharge data? A validation study. *Aust N Z J Obstet Gynaecol* 2008;48(5):481–4.
- [23] Yasmeen S, Romano PS, Schembri ME, Keyzer JM, Gilbert WM. Accuracy of obstetric diagnoses and procedures in hospital discharge data. *Am J Obstet Gynecol* 2006;194(4):992–1001.
- [24] Carre N, Uhry Z, Velten M, Tretarre B, Schwartz C, Molinie F, et al. Valeur prédictive et sensibilité du programme de médicalisation des systèmes d'information (PMSI) par rapport aux registres des cancers : application au cancer de la thyroïde (1999–2000). *Rev Epidemiol Sante Publique* 2006;54(4):367–76.
- [25] Remontet L, Mitton N, Couris CM, Iwaz J, Gomez F, Olive F, et al. Is it possible to estimate the incidence of breast cancer from medico-administrative databases? *Eur J Epidemiol* 2008;23(10):681–8.
- [26] Misset B, Nakache D, Vesin A, Darmon M, Garrouste-Orgeas M, Mourvillier B, et al. Reliability of diagnostic coding in intensive care patients. *Crit Care* 2008;12(4):R95.
- [27] Casez P, Labarere J, Sevestre MA, Haddouche M, Courtois X, Mercier S, et al. ICD-10 hospital discharge diagnosis codes were sensitive for identifying pulmonary embolism but not deep vein thrombosis. *J Clin Epidemiol* 2010;63(7):790–7.
- [28] Ministère de la Santé. Arrêté du 20 septembre 1994 relatif au traitement des données d'activité médicale et de coût. In: Journal Officiel de la République française, n° 242 du 18 octobre 1994. Paris; 1994: 14761.
- [29] Ministère de la Santé. Arrêté du 22 juillet 1996 relatif au recueil et au traitement des données d'activité médicale. In: Journal Officiel de la République française, n° 173 du 26 juillet 1996. Paris; 1996: 11308.
- [30] Ministère de la Santé. Loi n° 2003-1199 du 18 décembre 2003 de financement de la sécurité sociale pour 2004. In: Journal Officiel de la République française, n° 293 du 19 décembre 2003. Paris; 2003.
- [31] Ministère de la Santé et des Sports. ATIH : Guide méthodologique de production des informations relatives à l'activité médicale et à sa facturation en médecine, chirurgie, obstétrique et odontologie. In: Paris; 2010: 105.
- [32] Chantry AA, Deneux-Tharoux C, Cans C, Ego A, Quantin C, Bouvier-Colle MH. Hospital discharge data can be used for monitoring procedures and intensive care related to severe maternal morbidity. *J Clin Epidemiol* 2011;64(9):1014–22.
- [33] Comité national d'experts sur la mortalité maternelle : rapport du comité national d'experts sur la mortalité maternelle, 2001–2006. Paris: Inserm - InVS; 2010. p. 62.
- [34] Roberts CL, Bell JC, Ford JB, Hadfield RM, Algert CS, Morris JM. The accuracy of reporting of the hypertensive disorders of pregnancy in population health data. *Hypertens Pregnancy* 2008;27(3):285–97.
- [35] Klemmensen AK, Olsen SF, Wengel CM, Tabor A. Diagnostic criteria and reporting procedures for pre-eclampsia: a national survey among obstetrical departments in Denmark. *Eur J Obstet Gynecol Reprod Biol* 2005; 123(1):41–5.
- [36] Lydon-Rochelle MT, Holt VL, Nelson JC, Cardenas V, Gardella C, Easterling TR, et al. Accuracy of reporting maternal in-hospital diagnoses and intrapartum procedures in Washington State linked birth records. *Paediatr Perinat Epidemiol* 2005;19(6):460–71.
- [37] Calle JE, Saturno PJ, Parra P, Rodenas J, Perez MJ, Eustaquio FS, et al. Quality of the information contained in the minimum basic data set: results from an evaluation in eight hospitals. *Eur J Epidemiol* 2000;16(11):1073–80.
- [38] Thornton C, Makris A, Ogle R, Hennessy A. Generic obstetric database systems are unreliable for reporting the hypertensive disorders of pregnancy. *Aust N Z J Obstet Gynaecol* 2004;44(6):505–9.

- [39] Humphries KH, Rankin JM, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research: are clinical data derived from administrative databases a reliable alternative to chart review? *J Clin Epidemiol* 2000;53(4):343–9.
- [40] Romano PS, Yasmeen S, Schembri ME, Keyzer JM, Gilbert WM. Coding of perineal lacerations and other complications of obstetric care in hospital discharge data. *Obstet Gynecol* 2005;106(4):717–25.
- [41] Groupement pour la modernisation du système hospitalier : typologie des systèmes d'information des établissements de santé. Paris: GMSIH; 2008.
- [42] Dussaucy A, Viel JF, Mulin B, Euvrard J. L'outil PMSI : biais, sources d'erreurs et conséquences. *Rev Epidemiol Sante Publique* 1994;42(4):345–58.
- [43] White RH, Brickner LA, Scannell KA. ICD-9-CM codes poorly identified venous thromboembolism during pregnancy. *J Clin Epidemiol* 2004;57(9):985–8.
- [44] Geller SE, Ahmed S, Brown ML, Cox SM, Rosenberg D, Kilpatrick SJ. International Classification of Diseases-9th revision coding for preeclampsia: how accurate is it? *Am J Obstet Gynecol* 2004;190(6):1629–33 [discussion 1633–24].
- [45] Lombrail P, Minvielle E, Kohler F, Heve D, Mayeux D, Naiditch M, et al. Problèmes posés par le codage de l'information médicale dans le cadre du programme de médicalisation du système d'information hospitalier. *Rev Epidemiol Sante Publique* 1991;39(3):285–95.
- [46] Lydon-Rochelle MT, Holt VL, Cardenas V, Nelson JC, Easterling TR, Gardella C, et al. The reporting of pre-existing maternal medical conditions and complications of pregnancy on birth certificates and in hospital discharge data. *Am J Obstet Gynecol* 2005;193(1):125–34.
- [47] Demlo LK, Campbell PM, Brown SS. Reliability of information abstracted from patients' medical records. *Med Care* 1978;16(12):995–1005.
- [48] Demlo LK, Campbell PM. Improving hospital discharge data: lessons from the National Hospital Discharge Survey. *Med Care* 1981;19(10):1030–40.
- [49] Valleron AJ. L'épidémiologie humaine. Conditions de son développement en France et rôle des mathématiques. Académie de Médecine edn. Paris: EDP Sciences; 2006.