

## Estimation de l'incidence des gastro-entérites aiguës à partir des bases de données de l'Assurance maladie

*A method for assessing acute gastro-enteritis incidence from French health insurance databases*

Beaudeau P<sup>1</sup>, Bounoure F<sup>2</sup>, Genet M<sup>3</sup>

### Résumé

**Objectif** : L'Assurance maladie gère depuis 1998 une base de données des prestations de santé présentées au remboursement. Une étude a été menée pour créer un algorithme de discrimination des cas de gastro-entérites aiguës à partir des données de prescriptions médicamenteuses.

**Méthodes** : Quatre-vingt quatorze ordonnances contenant au moins un médicament utilisé pour le soin des gastro-entérites ont été collectées auprès de clients d'une pharmacie ainsi que le diagnostic associé. L'algorithme a été construit et testé sur cette base pour discriminer les cas de gastro-entérites aiguës des autres pathologies.

**Résultats** : La moitié de ordonnances échantillonnées étaient destinées au soin d'autres pathologies, comme les colopathies fonctionnelles, ou en prévention d'une diarrhée associée à une antibiothérapie. L'élaboration de l'algorithme de discrimination s'est basée essentiellement sur l'occurrence dans l'ordonnance de médicaments ou d'associations médicamenteuses spécifiques et sur la durée du traitement. L'algorithme produit un indicateur sensible et spécifique des gastro-entérites aiguës médicalisées. Quatre catégories de cas sont distinguées selon le nombre de médicaments prescrits et la présence d'anti-émétiques ou d'anti-infectieux intestinaux.

**Conclusion** : D'autres tests sont à prévoir pour évaluer le potentiel de ces indicateurs, notamment une validation croisée avec d'autres sources de données.

**Prat Organ Soins 2006;37(4): 313-23**

**Mots-clés** : médicament ; prescription médicale ; gastro-entérite ; surveillance épidémiologique ; France.

### Summary

**Aim**: The French health insurance set up in 1998 a drug prescription database in order to automatically process the refunding of insured patients for their health purchases. This study aimed to design an algorithm which can discriminate acute gastroenteritis cases from the drug prescription data.

**Methods**: We first collected a sample of 94 prescriptions including at least one drug indicated in acute gastroenteritis care. We also obtained the corresponding diagnosis by asking the patient. We then elaborated an algorithm for discriminating case and non case prescriptions.

**Results**: Half of the prescriptions were related to other diseases than acute gastroenteritis such as bowel chronic diseases or were preventive antidiarrhoeal prescriptions associated with antibiotherapy. The discriminating algorithm is mainly based upon the occurrence of non case specific drugs, case specific drug associations and the treatment duration. It results in a both sensitive and specific index for medicalized acute gastroenteritis incidence. The definition of case encompasses four sub-definitions to distinguish the cases according to the number of the prescribed drugs, and the occurrence of antiemetic or antibacterial drugs.

**Conclusion**: Further tests, including cross-validation, are needed to evaluate the potential of these indices.

**Prat Organ Soins 2006;37(4): 313-23**

**Keywords**: drug sales; France; gastroenteritis; physician's practice patterns; prescriptions; drug surveillance.

<sup>1</sup> Epidémiologiste, Institut de veille sanitaire, Saint-Maurice (France).

<sup>2</sup> Pharmacien, Laboratoire de pharmacie galénique, ADEN EA3234, Faculté de médecine pharmacie, Rouen (France).

<sup>3</sup> Pharmacien-conseil, chef de service, Direction régionale du service médical de Normandie, CNAMTS (France).

## INTRODUCTION

Les données de ventes de médicaments sont exploitées en épidémiologie pour estimer la prévalence de pathologies chroniques comme l'asthme [1, 2], le diabète [3-7], la sclérose en plaques [8], la maladie de Parkinson [9], l'épilepsie [10] ou les affections cardio-vasculaires [11]. Elles sont plus rarement utilisées pour estimer l'incidence de pathologies telles que les gastro-entérites aiguës [12], certaines pathologies respiratoires aiguës [13-15] ou les manifestations aiguës de l'allergie aux pollens [16]. Des méthodes spécifiques ont été proposées pour s'affranchir des incertitudes sur les posologies et sur l'observance des traitements [17, 18]. Celles-ci permettent donc d'estimer le nombre de cas, plutôt que de simplement compter les boîtes délivrées. En France, les bases de données de l'Assurance maladie (AM), contenant les informations relatives aux facturations des médicaments délivrés, limités aux médicaments prescrits et remboursés, autorisent le comptage direct des cas pathologiques [2, 7-10, 15, 19] ou l'étude des pratiques de prescription [20, 21].

L'objectif de notre étude était de proposer une procédure permettant de reconnaître les cas de gastro-entérites aiguës (GEA) médicalisées à partir des éléments de l'ordonnance présents dans les bases de données de l'AM. Cette procédure est destinée à l'exploitation ultérieure de ces bases pour la surveillance épidémiologique des GEA en France. Ce type d'approche a été initié et encouragé par le Service médical de l'AM [22].

Du fait de la fréquence élevée de leur incidence, les GEA représentent d'une part un coût important pour la collectivité dans les pays développés [23, 24] et d'autre part un indicateur sensible et réactif de la morbidité attribuable à l'exposition aux pathogènes d'origine entérique. La surveillance exercée en France par le réseau sentinelle des médecins généralistes est adaptée à l'étude des phénomènes épidémiologiques de grande ampleur comme les épidémies virales hivernales [25]. La sensibilité de ce système est cependant trop réduite (de l'ordre de 2/1000) pour déceler les phénomènes locaux et étudier leurs déterminants. Les données de l'AM, qui couvrent une part importante de la population française, pourraient pallier ce manque.

## MÉTHODES

### 1. Bases de données de l'Assurance maladie

#### a) Contenu

L'AM a développé dès les années 80 une base de données servant à la liquidation des actes de soins remboursés en

France. Le codage des médicaments y a été intégré en 1998. La version nommée ERASME (extraction recherches analyses pour un suivi médico-économique), développée en 2000, couvre le Régime général, y compris les sections locales mutualistes, soit 85,6 % de la population de France métropolitaine au 31 décembre 2002 (source Caisse nationale d'assurance maladie des travailleurs salariés CNAMTS). Ce taux connaît d'importantes variations locales (50-95 %) avec un minimum en secteur rural.

Depuis 2003, le système national d'information interrégimes de l'Assurance maladie (SNIR-AM), base anonymisée à vocation statistique, co-existe avec la base ERASME. Le code « bénéficiaire » anonymisé est identique à celui utilisé pour le Programme de médicalisation des systèmes d'information (PMSI), instauré dans les hôpitaux, ce qui assure la traçabilité de la consommation de soins des individus résidant en France.

Dans l'objectif de se prémunir d'une exploitation détournée des informations, les données sont effacées au bout de deux ou trois ans (respectivement pour ERASME et SNIR-AM).

Les principales variables renseignées dans ces bases de données et intéressant l'épidémiologie sont colligées en tableau I. Le terme « bénéficiaire » désigne le patient à qui la prescription est destinée ; il peut être l'assuré lui-même, mais aussi un des ayants droits qui lui sont rattachés (conjoint, enfants, personnes à charge). Comme l'adresse du bénéficiaire n'est pas présente dans les bases de données de l'AM, l'adresse utilisée pour la localisation de son lieu de résidence est celle de l'assuré.

Dans les bases de l'AM, les médicaments sont référencés par le code du Club inter pharmaceutique (CIP) qui correspond de façon univoque à une présentation commerciale d'un médicament. Les spécialités peuvent cependant être appelées par groupe homogène selon des nomenclatures chimiques et thérapeutiques : l'*Anatomical therapeutic classification* (ATC) de l'Organisation mondiale de la santé qui comprend cinq niveaux hiérarchiques jusqu'à la Dénomination commune internationale (DCI), ou la nomenclature de l'*European pharmaceutical marketing research association* (EPHmRA), moins détaillée. L'AM actualise la liste des CIP selon une fréquence hebdomadaire et met annuellement à jour la table de correspondance avec les nomenclatures ATC / EPHmRA.

#### b) Qualité des données

L'utilisation des données de l'AM est limitée par l'existence d'erreurs ou d'incertitudes (tableau I) affectant principalement l'adresse supposée du bénéficiaire. Ces erreurs proviennent soit de différences entre l'adresse du bénéficiaire et celle de l'assuré (étudiants rattachés à leurs parents, enfants de parents séparés inscrits au nom d'un parent mais domiciliés chez l'autre parent, etc.), soit d'erreurs sur l'adresse de l'assuré (changements d'adresse non

signalés, communication de l'adresse de travail ou de la mutuelle dans le cas des mutuelles décompteuses). Les mutuelles décompteuses comprennent les sections locales mutualistes (mutuelle générale de l'Education nationale, mutuelle de la Fonction publique, Mutuelle générale, etc.) et les mutuelles étudiantes. Elles fournissent à l'AM leur propre adresse au lieu de celle du bénéficiaire, dans la mesure où elles servent d'intermédiaires pour le remboursement à l'assuré de la

part revenant à l'AM. Elles protégeaient 12,7 % de la population française métropolitaine au 31 décembre 2002 (source : CNAMTS) et constituent la principale source d'erreurs sur la localisation du bénéficiaire. Par défaut de saisie, la date de prescription est remplacée par la date de délivrance. S'agissant des GEA, l'impact de ces erreurs reste limité : en 2000-2001, 90 % des malades achetaient leurs médicaments le jour même de la consultation [26].

Tableau I  
Contenu des bases de données de l'Assurance maladie sur les prescriptions médicamenteuses (ERASME, SNIIR-AM).

Nature de l'information	Qualité	Utilité
<i>Assuré</i>		
Matricule de sécurité sociale <sup>a</sup>	Obtention soumise à des conditions restrictives	
Matricule anonymisé FOIN <sup>b</sup>		Si suivi individualisé
Localisation de la résidence (adresse déclarée par l'assuré) :	DE = 10-20 %	Selon la résolution spatiale recherchée
Adresse complète <sup>a</sup>		
Code Postal (CP)		Intérêt si le code commune regroupe plusieurs CP (grandes villes)
Code INSEE de la commune	DM ≤ 2 % (province) DM = 5 % (rég. paris.)	Intérêt si le CP regroupe plusieurs communes (secteurs ruraux)
<i>Bénéficiaire</i>		
Sexe		
Date <sup>a</sup> ou année de naissance		
Rang dans la fratrie (naissances multiples) <sup>a</sup>		Si suivi individualisé
<i>Dossier</i>		
Code dossier ou archive <sup>a</sup>		Clé informatique
Références pharmacie dont CP		Repérage les achats de médicaments réalisés hors du secteur de résidence de l'assuré
Références médecin prescripteur dont CP		Idem
Date de délivrance		
Date de prescription	Moyennement fiable, saisie facultative	
<i>Produit</i>		
Code produit (CIP)		
Nom de spécialité		
Classe ATC		Appel par classe thérapeutique
Classe EPhMRA		Idem
Quantité délivrée (nombre de boîtes)		

<sup>a</sup> Information disponible seulement dans ERASME ou <sup>b</sup> dans SNIIR-AM.

DE : taux de données erronées.

DM : taux de données manquantes.

La grande majorité des dossiers de remboursement sont télétransmis par le pharmacien à la Caisse primaire d'assurance maladie (CPAM). La fréquence de télétransmission a été en moyenne de six par mois en 2003, mais elle variait de quotidienne à mensuelle selon les pharmacies. Dans moins de 10 % des cas, le client payait directement ses médicaments, puis envoyait par courrier sa feuille de soin à la CPAM pour se faire rembourser. Quatre-vingt-dix pour cent des feuilles de soin étaient saisies dans ERASME dans les deux mois qui suivent la délivrance des médicaments (Merlière Y (CNAMTS), communication personnelle le 26 juin 2003).

## 2. Méthodes

La définition de cas de GEA médicalisée fondée sur l'analyse des données de remboursement s'est appuyée sur l'observation des pratiques des médecins généralistes selon le protocole suivant :

- *Constitution de la liste des médicaments utilisés remboursés*

Les sources formelles d'information sur les médicaments indiqués (nomenclature ATC, références médicales opposables, conférences de consensus) reflètent mal la pratique réelle des médecins qui prescrivent de nombreux médicaments à service médical rendu considéré comme insuffisants, tels que les anti-infectieux intestinaux (classe A07A de l'ATC), les microorganismes antidiarrhéiques (A07F) et certains adsorbants intestinaux (A07B) [27]. Il a donc été jugé préférable d'établir la liste des médicaments utilisés dans le traitement des GEA (on mentionnera : médicaments *utilisés*) à dire d'experts de terrain, en l'occurrence de médecins généralistes et de pharmaciens d'officine [28].

- *Enquête en pharmacie*

L'enquête en pharmacie visait à constituer un échantillon d'ordonnances assorties du diagnostic associé. Elle a été menée par un étudiant en pharmacie dans le cadre de sa thèse d'exercice et s'est déroulée entre le 20 décembre 1999 et le 6 janvier 2000 dans une importante officine du centre du Havre. Le processus de recrutement a reproduit la procédure de sélection des dossiers informatiques dans les bases de données de

l'AM : tous les clients munis d'une ordonnance comportant au moins un médicament *utilisé* remboursé ont été sollicités pour être inclus dans l'étude. Le nom du médecin prescripteur, la date de consultation et de délivrance, le nom du médicament et le nombre de boîtes délivrées ont été tirés de l'ordonnance. Le client recruté a été interrogé sur l'âge et le sexe du malade, le code postal de son adresse de résidence, le diagnostic posé par le médecin prescripteur et la date d'apparition du premier symptôme (diarrhée ou vomissement) en cas de GEA.

- *Analyse discriminante*

L'analyse statistique de l'échantillon a permis de dégager de l'ordonnance les informations pertinentes pour distinguer les cas de GEA et les non-cas, puis de proposer un algorithme de discrimination. L'information utilisée s'est limitée à l'information transmise à l'AM afin d'assurer la pleine transposition de l'algorithme à l'exploitation des bases de données de l'AM. Les indicateurs de cas ont finalement été évalués en termes de sensibilité et de spécificité par rapport à la référence constituée par le diagnostic du médecin rapporté par le client.

La saisie et les calculs ont été faits sous Excel™.

L'échantillon d'ordonnances obtenu en pharmacie n'est pas un échantillon aléatoire de malades et de non-malades car le recrutement a été conditionné par la présence de médicaments *utilisés* dans l'ordonnance. Le calcul de sensibilité peut être considéré comme non biaisé si on admet que 100 % des cas de GEA qui ont consulté se sont vus prescrire au moins un médicament *utilisé*. Cette hypothèse paraît tout à fait raisonnable en regard des pratiques des généralistes décrites par le panel d'experts impliqués dans l'étude. L'échantillon des non-malades étant biaisé par construction, la spécificité est en revanche indéterminée et le calcul basé sur l'échantillon disponible produit une estimation sous-évaluée de la spécificité. La définition de cas triviale « présence d'un médicament *utilisé* dans l'ordonnance » conduit par exemple à une pseudo spécificité nulle. Si cette pseudo spécificité n'est pas conforme à la définition usuelle, elle constitue néanmoins une mesure relative de la spécificité et autorise la comparaison des différents indicateurs de cas.

La distribution de certaines variables telles que le nombre de classes thérapeutiques incluses dans les prescriptions ne correspond pas à des distributions classiques (en l'occurrence distribution de Poisson surdispersée). Pour tester la communauté des distributions de ces variables entre les cas et les non-cas, il a donc

fallu recourir à des méthodes non paramétriques. Le test proposé par Pitman [29] pour mettre en évidence une différence de valeur centrale (médiane, moyenne) entre deux échantillons a été mis en oeuvre.

## RÉSULTATS

Dans les sections qui suivent, on désigne conventionnellement par cas le cas de GEA diagnostiqué par le médecin prescripteur et rapporté par le client. Les non-cas recouvrent l'ensemble des autres diagnostics. Par opposition, l'écriture CAS et NON-CAS est réservée à la prédiction faite par l'algorithme de discrimination.

## 1. L'enquête en pharmacie

L'enquête en pharmacie a permis de collecter un échantillon de 94 ordonnances comportant au moins un médicament *utilisé*, soit 50 cas et 44 non-cas.

Les médicaments remboursés utilisés dans le soin des GEA se répartissent en six catégories principales (tableau II). Le nombre d'unités délivrées lors de l'enquête en pharmacie atteignait 234, dont 152 chez les cas, et se distribuait de façon très inégale entre les différentes classes thérapeutiques : antispasmodiques (29 %), produits microbiens (24 %), anti-émétiques (23 %), adsorbants intestinaux (11 %), ralentisseurs du transit (7 %) et anti-infectieux intestinaux (7 %).

Tableau II  
Principaux médicaments utilisés pour le traitement des gastro-entérites aiguës  
(Sources : Assurance maladie, Agence française de sécurité sanitaire des produits de santé).

Classes thérapeutiques	Classe ATC <sup>a</sup>	Principales spécialités	Principe actif (DCI)	Nombre spécialités <sup>b</sup>	Nombre d'unités remboursées <sup>c</sup>
<b>Antispasmodiques</b> intestinaux	(A03A)	Spasfon® Duspatalin® Débridat®	Phloroglucinol Mébévérine Trimébutine	43 (2)	275
<b>Anti-émétiques</b>	<b>A04A</b> A03F	Vogalène® Motilium® Peridys® Pimpéran®	Métopimazine Dompéridone Métoclopramide	60 (11)	207
<b>Produits microbiens</b> antidiarrhéiques	<b>A07F</b>  <b>A07B</b>	Ultralevure®  Lactéol® Carbolevure®	<i>Saccharomyces boulardii</i> Microorg. pdt acide lactique <i>S. cerevisiae</i>	45 (4)	146
<b>Ralentisseurs</b> du transit intestinal	<b>A07D</b>	Imodium® Arestal®	Lopéramide	108 (7)	80
Antisécrétoires	<b>A07X</b>	Tiorfan®	Acétorphan		
<b>Adsorbants</b> intestinaux	A07B Non classé A02X	Carbosylane® Bédélix® Smecta®	Charbon Montmorillonite Diosmectite	104 (3)	82
<b>Anti-infectieux</b> intestinaux	<b>A07A</b>	Ercéfuryl®	Nifuroxazide	115 (4)	34
<b>Préparations salines</b> pour réhydratation	Catégorie des dispositifs médicaux			7	0
<i>Total</i>				<i>482 (31)</i>	

<sup>a</sup> En gras : classe retenue comme cible dans son intégralité ; entre parenthèses : classe non retenue comme cible (voir texte).

<sup>b</sup> Entre parenthèses : nombre minimum de spécialités cumulant 90 % des ventes de la catégorie.

<sup>c</sup> Année 2000, France entière, nombre d'unités rapporté à 1000 habitants.

Les principales causes de prescription de médicaments utilisés à des patients atteints par d'autres pathologies que des GEA étaient la prévention des diarrhées consécutives à une antibiothérapie (n = 11, prescription de produits microbiens), les colopathies fonctionnelles (n = 6) et le reflux gastro-oesophagien (n = 4). Huit diagnostics de non-cas n'ont pas été identifiés précisément. La délivrance préventive pour le voyage, peu fréquente en janvier, n'a été observée qu'en vente conseil (ventes directes hors prescriptions médicales). Il n'a pas été observé de prescriptions de loperamide pour le soin des colopathies fonctionnelles, alors que, selon les médecins, cette pratique est répandue.

Les associations médicamenteuses prescrites dans le traitement des GEA incluaient une à quatre classes thérapeutiques. Les classes médicamenteuses utilisées en monothérapie (14 % des cas) étaient les produits microbiens et parfois les ralentisseurs du transit intestinal, à l'exclusion des autres classes thérapeutiques. La bithérapie (24 % des cas) associait souvent un anti-émétique, soit à un produit microbien, soit à un adsorbant. Les trithérapies (44 % des cas) associaient trois classes parmi les anti-émétiques, les adsorbants, les produits microbiens et les antispasmodiques. Les quadrithérapies (18 % des cas) délaissaient les produits microbiens et comportaient très souvent un antispasmodique, mais aussi un anti-infectieux, ce qui les distinguait des thérapies plus légères.

Les classes thérapeutiques utilisées dans le soin des GEA étaient représentées en plus grand nombre ( $p < 10^{-6}$ ) dans les ordonnances des cas (moyenne  $m = 2,68$  ; écart-type  $s = 0,94$ ) que dans celles des non-cas ( $m = 1,18$  ;  $s = 0,45$ ), tandis que le nombre de boîtes par classe thérapeutique utilisée dans le soin des GEA était plus important ( $p < 10^{-2}$ ) pour les non-cas ( $m = 1,65$  ;  $s = 1,04$ ) que pour les cas ( $m = 1,20$  ;  $s = 0,50$ ).

L'enquête en pharmacie a enfin montré que le décalage moyen entre l'apparition des symptômes chez les cas et la consultation était de 46 heures : 56 % des patients ont consulté dans les 24 heures et 74 % dans les 48 heures. L'achat des médicaments se faisait majoritairement le jour de la consultation, au plus tard le lendemain (délai moyen de dix heures).

## 2. Algorithme de discrimination et partition des CAS

L'algorithme de discrimination procède par exclusions successives de dossiers de remboursement, versés dans la catégorie des NON-CAS, puis opère une partition des CAS en fonction des médicaments présents. Les étapes, résumées en tableau III, sont les suivantes :

### *a) Exclusion des prescriptions pour lesquelles la délivrance est tardive*

La condition ((dates prescription - date délivrance) > deux jours) vise à éliminer les renouvellements d'ordonnance caractéristiques des pathologies chroniques et une partie des prescriptions faites à titre préventif, la recherche des produits étant alors moins empressée qu'en cas de diarrhée établie.

### *b) Exclusion des ordonnances comprenant un médicament éliminatoire*

Ces médicaments sont : méسالazine (Rowasa<sup>®</sup>, classe ATC A07E des anti-inflammatoires intestinaux), dédiée au soin des rectocolites hémorragiques et de la maladie de Crohn ; parmi les anti-infectieux intestinaux : (classe ATC : A07A), miconazole (Daktarin<sup>®</sup>), mystatine (Mycostatine<sup>®</sup>), amphotéricine B (Fungizone<sup>®</sup>), indiqués dans le traitement des candidoses, et tiliquinol, tilbroquinol (Intetrix<sup>®</sup>, classe ATC des antiprotozoaires P01), indiqués dans le traitement des amibiases qui sont principalement des maladies d'importation. La présence d'antibiotiques systémiques, généralement non indiqués dans le traitement des GEA, est aussi éliminatoire. La prescription conjointe d'un antidiarrhéique et d'un antibiotique vise le plus souvent à traiter à titre préventif une diarrhée iatrogène consécutive de l'antibiothérapie ; l'association Ultralevure<sup>®</sup> / Augmentin<sup>®</sup> est très représentative de cette situation et correspondait à environ la moitié des prescriptions d'Ultralevure<sup>®</sup> observées. Un petit nombre d'antibiotiques, comme la colistine et certains sulfamides (les spécialités contenant sulfaméthoxazole et triméthoprime), sont indiqués pour le traitement des GEA d'étiologie bactérienne et ne provoquent pas l'exclusion. Les fluoroquinolones sont aussi actives sur les bactéries entériques pathogènes mais elles ne sont pratiquement utilisées que pour des infections urinaires ou respiratoires. Dans cette famille, seul le Ciflox<sup>®</sup> est prescrit en cas de GEA mais, s'agissant alors majoritairement de cas importés ou de délivrances préventives en vue de voyages, l'exclusion est maintenue.

### *c) Exclusion des ordonnances correspondant à des traitements de longue durée*

La durée de traitement de première intention d'une GEA est volontairement limitée à huit à dix jours par le prescripteur. L'exclusion des ordonnances correspondant à une durée de traitement supérieure vise notamment à éliminer les ordonnances destinées au traitement de la maladie de Crohn, repérables à la prescription de quantités importantes de loperamide. Deux présentations d'un même médicament peuvent être interprétées différemment. Ainsi, la prescription de Smecta 90 doses, correspondant à 30 jours de traitement, évoque le traitement symptomatique des douleurs liées aux colopathies fonc-



Tableau III  
Définitions<sup>a</sup> simplifiées de CAS de gastro-entérites aiguës médicalisée  
basées sur les délivrances de médicaments remboursés.

CAS MONO	(Un anti-diarrhéique, i.e. ralentisseur du transit ou produit microbien, à l'exclusion d'autre CTU <sup>b</sup> ) et (Nombre de boîtes par CTU < 3).
CAS SEPT	(Au moins 2 CTU représentées dont un anti-infectieux intestinal) et (Nombre de boîtes par CTU < 4).
CAS EMET	(Non CAS SEPT) et (Plus de 2 CTU représentées dont un anti-émétique) et (Nombre de boîtes par CTU < 3).
CAS AUTRE	(Non CAS SEPT) et (Non CAS EMET) et (Au moins 2 CTU représentées) et (Nombre de boîtes par CTU < 3).
CAS	CAS MONO ou CAS SEPT ou CAS EMET ou CAS AUTRE
NON-CAS	(Non CAS) et au moins une CTU représentée

<sup>a</sup> Les règles d'exclusion basées sur la présence de médicaments non indiqués sont données dans le texte de l'article.

<sup>b</sup> CTU : classe thérapeutique *utilisée* dans le traitement des GEA (définition en tableau II).

tionnelles ou aux gastrites tandis que la prescription de Smecta 30 doses est typique des GEA.

**d) La sélection des CAS prend enfin conjointement en compte le nombre de médicaments utilisés présents dans l'ordonnance, leur nature et le nombre de boîtes délivrées par médicament**

Sont considérées comme CAS : (i) les ordonnances qui ne comportent qu'un médicament *utilisé* (monothérapie) à condition que le médicament soit un ralentisseur du transit ou un produit microbien, et que le nombre d'unités délivrées corresponde à une durée de traitement d'une dizaine de jour au plus ; (ii) les bithérapies (deux classes de médicaments utilisés représentées dans l'ordonnance) si le nombre d'unités délivrées par classe est inférieur ou égal à deux ; (iii) les thérapies d'ordre supérieur si le nombre unités délivrées par classe est inférieur ou égal à trois.

Du fait de la diversité des prescriptions dédiées au traitement des GEA, une partition des CAS a été opérée. Cette catégorisation repose d'une part sur le nombre des médicaments prescrits et d'autre part sur la présence dans l'ordonnance de classes thérapeutiques particulières. CAS MONO correspond à une monothérapie. Concernant les multithérapies, CAS SEPT signe la présence d'un anti-infectieux et CAS EMET celle d'un anti-émétique dans l'ordonnance. Si les deux médicaments sont présents le CAS est versé dans la catégorie CAS SEPT.

L'algorithme a été implémenté sous la forme d'un logiciel développé en langage C. Ce programme admet en entrée les fichiers fournis par l'AM. Il génère en sortie des séries chronologiques du nombre de CAS quotidien détaillé par catégorie.

### 3. Evaluation des indicateurs obtenus

La définition de CAS triviale « présence d'un médicament *utilisé* dans l'ordonnance », possède de par le mode d'échantillonnage, une sensibilité de 100 % et une spécificité nulle. L'exclusion des antispasmodiques a laissé la sensibilité inchangée et a accru médiocrement la spécificité (30 %). Quelle que soit la classe thérapeutique *utilisée*, sa présence dans l'ordonnance n'indiquait pas la pathologie à la fois avec sensibilité et avec spécificité (tableau IV). Une définition de CAS conforme à l'algorithme a permis en revanche d'obtenir une bonne spécificité (86 %) sans altérer fortement sa sensibilité (88 %). Les six NON-CAS classés par erreur en CAS se répartissaient entre CAS SEPT (3/27) et CAS MONO (3/8), cette catégorie ayant été la moins spécifique.

### 4. Préconisations pour la requête d'extraction

La requête d'extraction de la base de l'AM doit spécifier (i) les critères de sélection des ordonnances et (ii) la nature des informations à extraire.

La population étudiée peut être une population domiciliée ou une population résidant (population de passage incluse) sur un territoire donné pendant une période définie<sup>1</sup>. Dans le premier cas, le test pour l'extraction porte alors sur le code postal de l'assuré et la date de prescription ; dans l'autre cas, le code postal testé est celui du médecin prescripteur ou de la pharmacie qui a délivré les médicaments. Sauf cas particulier (constitution de statistiques agrégées à l'échelle nationale, etc.), les dossiers instruits par les mutuelles *décomp-teuses* sont à exclure.

Les médicaments cibles, c'est-à-dire les médicaments dont la rencontre dans les fichiers provoque l'extraction

<sup>1</sup> La population de passage n'est pas accessible au niveau local (CPAM) mais seulement au niveau national.

de l'ordonnance et des informations associées (tableau I), correspondent aux médicaments *utilisés* (tableau II), à l'exception des antispasmodiques trop peu spécifiques du traitement des GEA.

Dans la requête d'extraction, les médicaments cibles des GEA sont à libeller soit par classe ATC (ou EPhMRA) quand l'ensemble des médicaments de la classe est pertinent, soit par CIP quand seules quelques spécialités de la classe d'appartenance présentent un intérêt. Quatre classes thérapeutiques (produits microbiens, ralentisseurs et antisécrétoires, anti-inflammatoires et anti-infectieux intestinaux) correspondent à des subdivisions de la classe ATC A07 des « antidiarrhéiques, anti-inflammatoires et anti-infectieux intestinaux ». D'autres classes thérapeutiques sont à recomposer : le métopramide (Primpéran) et le dompéridone (Motilium), couramment utilisés pour leurs propriétés anti-émétiques, sont en effet classés dans la catégorie des stimulants de la motricité intestinale (A03F) et l'argile diosmectite (Smecta) très utilisée comme adsorbant intestinal pour le traitement des GEA, est classée comme anti-acide par EPhMRA (A02).

Les 31 spécialités les plus vendues, soit 16 % des spécialités « utilisées » remboursées, cumulaient 90 % du volume des délivrances en 2000 (tableau II). Cela encourage l'emploi de listes simplifiées en situation d'investigation rapide, au prix d'une perte d'information négligeable (tableau III).

L'extraction de la totalité des lignes de facturation, plutôt que des seules lignes correspondant aux médicaments cibles, permet d'améliorer la spécificité de l'indicateur en introduisant des règles d'exclusion fondées sur la présence de médicaments non cibles, voire non indiqués, comme la plupart des antibiotiques systémiques.

## DISCUSSION

L'enquête en pharmacie a montré qu'une analyse méthodique du contenu des prescriptions permet de distinguer efficacement les cas de GEA parmi les autres diagnostics possibles. Cette approche s'est montrée beaucoup plus performante que le simple décompte du nombre de boîtes remboursées (nombre de cas de GEA médicalisés sous-jacents inconnu) et que le comptage des occurrences de classes thérapeutiques *utilisées* dans les prescriptions présentées au remboursement (sensibilité ou bien spécificité déficiente selon la classe). En effet, à l'exception des anti-infectieux intestinaux, aucune catégorie de médicaments n'est spécifique des GEA, car il s'agit de traitements de symptômes eux-mêmes non spécifiques des GEA. C'est pourquoi il a été nécessaire de fonder l'algorithme sur des règles concernant la nature des associations entre classes thérapeutiques présentes et les quantités délivrées.

L'application de cette méthode à l'exploitation des

Tableau IV  
Sensibilité et spécificité (avec leur intervalle de confiance à 95 %) des classes thérapeutiques dans une ordonnance présentée au pharmacien comme indicateur de cas médicalisés de gastro-entérite aiguë (N = 94 ordonnances).

Classe thérapeutique	Sensibilité		Spécificité	
Antispasmodique	54 %	[40 ; 68]	57 %	[42 ; 71]
Ralentisseur du transit	30 %	[17 ; 43]	98 %	[93 ; 100]
Produit microbien	48 %	[34 ; 62]	73 %	[60 ; 86]
Adsorbant	38 %	[25 ; 51]	91 %	[82 ; 99]
Anti-émétique	62 %	[49 ; 75]	68 %	[54 ; 82]
Anti-infectieux	30 %	[17 ; 43]	98 %	[93 ; 100]
Une classe au moins	100 %	-	0 %	-
Une classe au moins sauf antispasmodique	100 %	-	30 %	[16 ; 43]
CAS (selon algorithme)	88 %	[79 ; 97]	86 %	[76 ; 97]



bases de données de l'AM pose cependant des problèmes de localisation spatiale des CAS et d'adaptation de la définition de CAS aux changements contextuels.

### 1. Localisation du bénéficiaire

Les erreurs sur le lieu de résidence du bénéficiaire contenues dans les bases de l'AM peuvent être mitigées en excluant de l'analyse les mutuelles *décompteuses*. Cette exclusion possède cependant une contrepartie :

- Impossibilité de standardiser les taux d'incidence sur les caractères socio-démographiques usuels. La structure en âge de la population protégée par les mutuelles *décompteuses* n'est pas connue ni assimilable à celle de la population générale, notamment du fait de la surreprésentation des jeunes adultes (les mutuelles étudiantes couvriraient en 2002 2,6 % de la population française) ;
- Impossibilité d'estimer précisément le taux d'incidence brut dès qu'on utilise les données à une échelle infra-départementale car il n'existe pas de données démographiques par régime d'assurance maladie pour des unités géographiques plus fines que le département. Ce problème sera résolu quand les mutuelles *décompteuses* livreront à l'AM l'adresse de résidence de leurs assurés. Les erreurs concernant l'adresse de résidence du bénéficiaire se réduiront alors aux cas des ayants droits n'habitant pas à la même adresse que l'assuré, c'est-à-dire essentiellement des enfants et des jeunes adultes.

Enfin, si l'étude vise à connaître l'effet des facteurs de risque, il s'agit de localiser le malade au moment de son exposition potentielle. Ceci n'est pas possible de façon générale, mais il est toutefois envisageable de réduire les erreurs liées aux déplacements de longue durée. A cet effet, on peut soustraire de la population d'étude les dossiers pour lesquels la localisation du médecin prescripteur, ou de la pharmacie qui a délivré les médicaments, diffère de celle de l'assuré (par exemple, en testant l'égalité des codes de département). Concernant les déplacements quotidiens, l'épidémiologiste ne peut que restreindre sa population d'étude aux catégories les plus sédentaires comme les enfants ou les personnes âgées.

### 2. Adaptation de l'algorithme à l'évolution du contexte

L'homogénéité dans le temps des indicateurs de CAS est compromise par tous les changements affectant la pharmacopée (apparition de nouvelles molécules), la thérapeutique, le commerce (renouvellement des spé-

cialités, apparition de génériques) ou la prise en charge au titre de la sécurité sociale. Pour les collectes d'informations s'inscrivant dans la durée, l'utilisateur se doit de veiller à la préservation de cette homogénéité, d'une part en mettant régulièrement à jour la liste des codes médicaments, et d'autre part en adaptant l'algorithme de discrimination aux modifications du contexte.

Lors de l'écriture de la requête, la désignation des médicaments cibles par classe thérapeutique (ATC ou EPhMRA) peut être préférée à leur désignation par spécialité commerciale (CIP) car ces nomenclatures chimiques sont indépendantes des mouvements commerciaux et actualisées annuellement. Cependant, le délai d'intégration des CIP dans ces nomenclatures cause une perte d'information dont l'importance dépend du taux de renouvellement des spécialités, très variable selon la classe thérapeutique et plutôt faible en ce qui concerne les spécialités utilisées dans le traitement de GEA.

Les changements de pratique de prescription peuvent être consécutifs à de nouvelles recommandations thérapeutiques (par exemple la promotion de la réhydratation orale pour le traitement de la diarrhée chez les enfants) ou à des changements de prise en charge. L'adaptation de la définition de CAS aux changements de pratique paraît plus délicate que celle à l'apparition de nouvelles spécialités car il faut alors en revoir les règles de discrimination. Concernant l'épidémiologie des GEA, le déremboursement partiel ou total des médicaments à service médical rendu insuffisant apparaît comme la perspective la plus préoccupante. Le déremboursement des produits microbiens programmé en 2006 entraînera non seulement une réorientation des prescriptions mais aussi la disparition de ces spécialités dans les bases de l'AM.

### 3. Perspectives

En l'absence de données collectées spécifiquement pour le test des indicateurs, les performances de l'indicateur CAS ont été calculées sur les données ayant servi à l'ajustement de l'algorithme de discrimination, ce qui conduit à leur surévaluation.

De plus, les données ont été collectées dans un contexte restreint (94 ordonnances présentées dans une pharmacie du Havre en hiver 2000-2001) susceptible de biaiser le choix des règles de discrimination. La répartition par classe thérapeutique des boîtes délivrées lors de l'enquête en pharmacie diffère en effet ( $p = 0,02$ ) de celle des médicaments remboursés par l'AM en 2000. Le rejet de l'hypothèse nulle<sup>2</sup> s'explique notamment par le fait que les anti-infectieux intestinaux et les produits microbiens ont été proportionnellement moins représentés au

<sup>2</sup> L'hypothèse nulle est l'hypothèse statistique que le chercheur remet en question en montrant que les résultats de l'échantillon ne sont pas compatibles avec elle.

plan national sur l'ensemble de l'année que dans l'enquête (respectivement - 41 % et - 25 %).

Pour ces raisons, il serait intéressant de reproduire l'étude dans d'autres villes et sur l'ensemble des saisons pour mieux évaluer l'algorithme de discrimination utilisé, et au besoin l'amender.

Un échantillon plus important d'ordonnances associées aux diagnostics permettrait aussi d'envisager la construction d'un algorithme spécifique dédié aux jeunes enfants, particulièrement touchés par la pathologie étudiée. Certaines formes pathologiques, comme le reflux gastro-oesophagien ou de la conjonction d'une GEA et d'une infection non entérique, sont plus fréquentes durant l'enfance qu'aux autres âges. Il peut en résulter des prescriptions (anti-émétiques ou antibiotiques, respectivement) que l'algorithme interprète mal dans sa consistance actuelle. Les prescriptions pédiatriques sont aussi plus parcimonieuses que les prescriptions pour adultes et certaines prescriptions pédiatriques d'anti-émétiques ou d'antispasmodiques en monothérapie seraient ainsi actuellement exclues à tort de la définition de CAS. Enfin la promotion récente du traitement par réhydratation orale pour le traitement de la diarrhée chez les enfants de moins de cinq ans justifie l'introduction d'une règle spécifique, dans la mesure où ces produits sont désormais remboursés (juin 2003) et codés (codage LPP 1175983 de la Liste des produits et des prestations) dans les bases de données de l'AM (Merlière Y (CNAMTS), communication personnelle, le 8 décembre 2005).

## CONCLUSION

Un échantillon d'ordonnances couplées au diagnostic du prescripteur a été collecté en pharmacie. L'exploitation de ces données a montré qu'il était possible de reconnaître les cas de GEA sur la base de l'information contenue dans l'ordonnance, même limitée aux données de facturation relatives aux médicaments remboursés.

L'algorithme de discrimination obtenu permet ainsi d'exploiter les bases de données de l'AM pour l'étude épidémiologique des GEA. Le potentiel de résolution spatio-temporelle des données de l'AM est très supérieur à celui des autres systèmes d'information dédiés à la surveillance des GEA dans la mesure où elles sont disponibles par commune et par jour sur l'ensemble du territoire français et quasi exhaustives (SNIIR-AM). Le problème résiduel concerne la localisation de la résidence d'une fraction des cas (environ 20 %).

Le contenu des bases de données de l'AM est cependant susceptible d'évolution en cas de changements dans la pharmacopée, dans les habitudes de prescription des médecins ou dans les règles de remboursement. Ce risque potentiel réclame de la vigilance de la part de l'utilisateur et nécessitera une révision de l'algorithme de discrimination des cas de GEA si un changement important survient dans ce contexte.

L'algorithme est en cours d'évaluation sur un large échantillon de données fourni par l'AM, notamment dans des contextes documentés au plan épidémiologique.

## Remerciements

*Nous remercions Mme Gauffriau et M. Merlière de la Caisse nationale de l'assurance maladie des travailleurs salariés (CNAMTS) qui nous ont conseillé en vue d'exploiter les bases de données de l'Assurance maladie. Nous remercions également Mme Rambourg-Schepens pour son aide précieuse.*

## RÉFÉRENCES

1. Kesten S, Rebuck AS, Chapman KR. Trends in asthma and chronic obstructive pulmonary disease therapy in Canada, 1985 to 1990. *J Allergy Clin Immunol* 1993; 92:499-506.
2. Deprez P, Chinaud F, Clech S, Vallier N, Germanaud J, Weill A et al. La population traitée par la classe des antiasthmatiques en France métropolitaine : les données du régime général de l'Assurance maladie, 2000. *Rev Med Ass Maladie* 2004; 35:3-11.
3. Fontbonne A, Papoz L, Eschwege E. Drug sales data and prevalence of diabetes in France. *Rev Epidemiol Sante Publique* 1986;34:100-5.
4. Papoz L. Utilization of drug sales data for the epidemiology of chronic diseases: the example of diabetes. *The EURODIAB Subarea C Study Group. Epidemiology* 1993;4:421-7.
5. Vauzelle-Kervroedan F, Forhan A, Papoz L. Regional prevalence of diabetes treated with oral hypoglycemic agents. *Diabete Metab* 1993;19:291-5.
6. Sartor F, Walckiers D. Estimate of disease prevalence using drug consumption data. *Am J Epidemiol* 1995;141:782-7.
7. Ricordeau P, Weill A, Vallier N, Bourrel R, Guilhot J, Fender P,

- et al. *Prévalence et coût du diabète en France métropolitaine : quelles évolutions entre 1998 et 2000 ? Rev Med Ass Maladie* 2002; 33:257-65.
8. Kazaz ES, Grisouard S, Zanni JL, Lemoine G, Weill A, Vallier N, Fender P, et al. *La sclérose en plaque traitée par interféron b : prévalence et coût pour l'Assurance maladie en 2000. Rev Med Ass Maladie* 2003; 34:147-56.
9. Lajugie D, Bertin N, Chantelou ML, Vallier N, Weill A, Fender P, et al. *Prévalence de la maladie de Parkinson et coût pour l'Assurance maladie en 2000 en France métropolitaine. Rev Med Ass Maladie* 2005; 36:113-22.
10. Karsenty D, Bourdel H, Cherrier C, Dumas Y, Vallier N, Weill A, et al. *La population traitée par médicaments antiépileptiques en France métropolitaine : les données du régime général de l'Assurance maladie. Rev Med Ass Maladie* 2005; 36:123-32.
11. Oreberg M, Jonsson GG, West K, Eberhard-Grahn M, Rastam L, Melander A. *Large intercommunity difference in cardiovascular drug consumption: relation to mortality, risk factors and socioeconomic differences. Eur J Clin Pharmacol* 1992; 43:449-54.
12. Beaudeau P, Payment P, Bourderont D, Mansotte F, Boudhabay O, Laubies B, et al. *A time series study of anti-diarrheal drug sales and tap-water quality. Int J Environ Health Res* 1999; 9:293-311.
13. Zeghnoun A, Beaudeau P, Carrat F, Delmas V, Boudhabay O, Gayon F, et al. *Air pollution and respiratory drug sales in the city of Le Havre, France, 1993-1996. Environmental Research* 1999; 224-30.
14. Pitard A, Zeghnoun A, Courseaux A, Lamberty J, Delmas V, Fossard JL, et al. *Short-term associations between air pollution and respiratory drug sales. Environ Res* 2004; 95:43-52.
15. Allenbach D, Montagnier B, Souche A, Vallier N, Weill A, Chinaud F, et al. *La population traitée par médicaments antituberculeux en 2003 : les données du régime général de l'Assurance maladie. Rev Med Ass Maladie* 2004; 35:223-32.
16. Harf R, Dechamp C. *Pollinosis and use of anti-allergic drugs: ragweed in the Rhone-Alpine region. Rev Mal Respir* 2001; 18:517-22.
17. Spitzer WO, Suissa S, Ernst P, Horwitz RI, Habbick B, Cockcroft D, et al. *The use of beta-agonists and the risk of death and near death from asthma. N Engl J Med* 1992; 326:501-6.
18. Osborne ML, Vollmer WM, Johnson RE, Buist AS. *Use of an automated prescription database to identify individuals with asthma. J Clin Epidemiol* 1995; 48:1393-7.
19. Claroux-Bellocq D, de Baillencourt S, Saint-Jean F, Chinaud F, Vallier N, Weill A, et al. *Les traitements de substitution aux opiacées en France métropolitaine en 2000 : les données du régime général de l'Assurance maladie. Rev Med Ass Maladie* 2003; 34:93-102.
20. Lecadet J, Vidal P, Baris B, Vallier N, Fender P, Allemand H. *Médicaments psychotropes: consommation et pratiques de prescription en France métropolitaine.I. Données nationales, 2000. Rev Med Ass Maladie* 2003; 34:75-84.
21. Lecadet J, Vidal P, Baris B, Vallier N, Fender P, Allemand H. *Médicaments psychotropes: consommation et pratiques de prescription en France métropolitaine.II. Données et comparaisons régionales, 2000. Rev Med Ass Maladie* 2003; 34:233-48.
22. Fender P, Weill A. *Epidémiologie, santé publique et bases de données médico-tarifaires. Rev Epidemiol Santé Publique* 2004; 52:113-7.
23. Payment P. *Epidemiology of endemic gastrointestinal and respiratory diseases : incidence, fraction attributable to tap water and costs to society. Wat Sci Tech* 1997; 35:11-2.
24. Fourquet F, Desenclos JC, Maurage C, Baron S. *Acute gastroenteritis in children in France: estimates of disease burden through national hospital discharge data. Arch Pediatr* 2003; 10:861-8.
25. Desenclos JC, Rebiere I, Letrillard L, Flahault A, Hubert B. *Diarrhoea-related morbidity and rotavirus infection in France. Acta Paediatr Suppl* 1999; 88:42-7.
26. Bounoure F. *Utilisation des médicaments comme marqueur de surveillance épidémiologique des gastro-entérites aiguës. Thèse de Pharmacie. Faculté mixte de médecine et de pharmacie, Rouen, 2002.*
27. Naudin F, Sermet C. *La prescription de médicaments à service médical rendu insuffisant an 2001. Institut de Recherche et Documentation en Economie de la Santé, editor: 1-151. 2004. Paris, IRDES. Série Analyse.*
28. Beaudeau P, Bonmati C, Boudhabay O, Bourderont D, Delmas V, Gayon F, et al. *De l'usage de la consommation médicamenteuse pour la surveillance des effets sanitaires de la pollution atmosphérique au Havre. Pollution atmosphérique* 1994; 3:133-42.
29. Pitman EJG. *Significance tests which may be applied to samples from any populations. Suppl Journ Royal Stat Soc* 1937; 4:119-30.