

UNIVERSITE PARIS-DESCARTES
ECOLE DOCTORALE 420 SANTE PUBLIQUE

Année 2015

Thèse de Doctorat en Santé Publique
Spécialité : Epidémiologie

Mathilde PIVETTE

**Surveillance des maladies infectieuses à partir des
ventes de médicaments en pharmacies**

Thèse dirigée par :

Pr. Avner BAR-HEN

Dr. Judith MUELLER

Jury :

M. Pierre-Yves BOELLE

Mme Annie FOURRIER-REGLAT

M. Pascal ASTAGNEAU

M. Loïc JOSSERAN

M. Avner BAR-HEN

Mme Judith MUELLER

Rapporteur

Rapporteur

Examineur

Examineur

Directeur de thèse

Co-directeur de thèse

« Thèse préparée dans le cadre du Réseau doctoral en santé publique animé par l'EHESP »

REMERCIEMENTS

Je remercie Monsieur Pierre-Yves BOELLE et Madame Annie FOURRIER-REGLAT d'avoir accepté d'être rapporteurs de cette thèse et pour leur implication dans l'évaluation de ce travail.

Je remercie Monsieur Pascal ASTAGNEAU et Monsieur Loïc JOSSERAN pour avoir accepté d'être examinateurs de cette thèse.

Je tiens à remercier vivement mes directeurs de thèse, Monsieur Avner BAR-HEN et Madame Judith Mueller, pour m'avoir guidée et soutenue dans mes travaux de recherche pendant ces trois dernières années. Leurs conseils et leur disponibilité m'ont permis d'aller au bout de ce travail de recherche avec succès, malgré les obstacles.

Je remercie la société Celtipharm pour le financement de la thèse via une convention CIFRE.

Merci à Pascal Crépey pour son implication dans la mise en place du projet de recherche et ses conseils dans la réalisation de cette thèse.

Je remercie l'équipe de département EpiBiostat, qui m'a permis de réaliser cette thèse dans une ambiance de travail très agréable et stimulante. Merci à la direction du Réseau Doctoral de l'EHESP pour le soutien logistique pendant cette thèse. Je remercie les doctorants du Réseau avec qui j'ai pu échanger, notamment Anthony et Michael avec qui j'ai pris plaisir à administrer l'association Ph'Doc.

Je remercie les personnes qui ont partagé mon bureau pendant quelques mois ou quelques années pour leur soutien et leur bonne humeur, notamment Ilana, Christophe et Mathilde.

Je remercie ma famille: mes parents, mes frères, Simon et Jean-Marie, Lucie et Yann qui m'encouragent dans mes projets. Merci à Pascal pour son précieux soutien.

RESUME

Le suivi des ventes de médicaments en pharmacies est un outil de surveillance qui s'est développé au début des années 2000 pour la surveillance des maladies infectieuses et la détection d'épidémies. Ces données présentent en effet de nombreux avantages pour la surveillance de par le volume important, l'exhaustivité et la rapidité d'obtention des données collectées de façon automatique. L'objectif de cette thèse était de déterminer l'intérêt du suivi des ventes de médicaments pour la surveillance et la prévention des maladies infectieuses en France. Les données de ventes analysées dans cette thèse sont issues d'un échantillon de 3004 pharmacies en France Métropolitaine, développé par la société *Celtipharm*. Trois axes de recherche ont été développés dans cette thèse :

Nous avons tout d'abord réalisé une revue systématique de la littérature sur la surveillance des maladies infectieuses à partir des ventes de médicaments. Vingt-sept articles ont été analysés et les résultats montrent que le suivi des ventes de médicaments est un outil valide pour la surveillance et la détection de la gastro-entérite et de la grippe. De plus, les médicaments achetés sans ordonnance ont le potentiel d'émettre une alerte plus précoce que les systèmes de surveillance traditionnels. Des études pourraient être développées pour la surveillance d'autres maladies infectieuses.

Notre deuxième étude avait pour objectif d'évaluer l'intérêt du suivi des ventes de médicaments pour la détection des épidémies saisonnières de gastro-entérites en France. Les ventes de médicaments indiqués en cas de gastro-entérite ont été comparées au nombre de cas vus en consultation entre 2008 et 2013. La méthode de *Serfling* a été utilisée pour détecter les périodes épidémiques. Les cinq épidémies saisonnières ont pu être détectées à partir des ventes de médicaments, sans fausse alerte. De plus, les médicaments non-prescrits ont permis une détection précoce des épidémies, en moyenne 2,25 semaines plus tôt que les données de référence.

La troisième étude a porté sur l'estimation de la couverture vaccinale (CV) contre la grippe et le rotavirus à partir des ventes de vaccins en pharmacies. Nous avons pu estimer la couverture vaccinale grippe chez l'ensemble de la population, et avons développé une méthode pour estimer en temps-réel la CV dans la population cible des plus de 65 ans. Il n'existe actuellement pas d'estimation de la CV rotavirus en France car le vaccin n'est pas remboursé, et nos analyses ont permis d'apporter une information sur cette couverture depuis 2008.

Les études développées dans cette thèse ont permis de démontrer que le suivi des ventes de médicaments était un outil valide et utile pour la surveillance des maladies infectieuses. Ses intérêts principaux résident dans la rapidité d'obtention des données qui offre la possibilité de déclencher rapidement une alerte et l'obtention de données de ventes de médicaments non-remboursés qui permettent d'avoir des informations sur des pathologies légères ou les premiers symptômes de maladies.

SUMMARY

Drug sales data analysis is a tool that has been developed in the early 2000s for infectious diseases surveillance and outbreak detection. These data present indeed many advantages for disease surveillance: the large volume and the completeness of the data, the rapidity to obtain the data automatically collected from pharmacies. The objective of this thesis was to determine the value of drug sales analysis for infectious disease surveillance in France. Sales data analyzed in this thesis were collected from a sample of 3,004 pharmacies in Metropolitan France, set up by the company *Celtipharm*. Three research axes have been developed in this thesis:

We first did a systematic literature review on the surveillance of infectious diseases based on drug sales data. Twenty-seven articles have been analyzed and results showed that drug sales data analysis is a valid tool for the surveillance and outbreak detection of gastroenteritis and influenza. Besides, over-the-counter (OTC) drugs have the potential to detect outbreaks earlier than traditional surveillance data. Further studies could be developed for other infectious diseases.

The second study focused on the detection of seasonal gastrointestinal disease outbreak in France based on drug sales data. Drug sales data have been compared to the number of cases reported by a network of GPs between 2008 and 2013. The *Serfling* method was used to detect epidemic periods in the times series. The five seasonal epidemic periods have been detected based on drug sales data, without false alert. Moreover, non-prescribed drugs allowed the detection of outbreaks on average 2.25 weeks earlier than reference data.

The third study focused on the estimation of vaccination coverage (VC) against influenza and rotavirus using vaccine sales in pharmacies. We estimated the influenza vaccination coverage in the whole population and developed a method to estimate the VC in real time in the target population of the ≥ 65 years old. There is no estimation of rotavirus vaccination coverage in France as the vaccine is not reimbursed and we were able to estimate vaccination coverage since 2008 based on vaccine sales.

The studies developed in this thesis showed that drug sales data analysis is a valid and useful tool for infectious disease surveillance. The main interests of this method are the rapidity to obtain the data that allows early outbreak detection and the collection of non-prescribed drugs that can reflect mild diseases or the first symptoms of a disease.

Mots-clés:

Surveillance; Ventes de médicaments; Maladies infectieuses; Détection d'épidémies; Pharmacies; France Métropolitaine

Key Words:

Surveillance; Drug sales; Infectious diseases; Outbreak detection; Pharmacies; Metropolitan France

SOMMAIRE

Liste des publications et communications	8
Liste des abréviations.....	10
I. Préambule.....	11
II. Introduction	12
1) La surveillance épidémiologique.....	12
a) Définition.....	12
b) Surveillance directe	13
c) Surveillance syndromique	13
d) Surveillance à partir des ventes de médicaments	15
2) La base de données <i>Celtipharm</i> et le médicament en France.....	17
a) Le panel Xpr-SO®.....	17
b) Utilisation de la base de données pour des études épidémiologiques	18
c) Les pharmacies en France.....	19
d) Le marché des produits officinaux en France.....	19
(i) Le médicament : définition et classification	19
(ii) Les chiffres du marché officinal du médicament.....	21
(iii) Autres produits de santé.....	22
e) La classification des médicaments.....	22
3) Problématique scientifique	23
III. Revue de la littérature sur la surveillance des maladies infectieuses à partir des ventes de médicaments.....	26
1) Résumé.....	26
2) Article.....	28
IV. Surveillance de la gastro-entérite en France à partir des ventes de médicaments ...	43
1) Résumé.....	43
2) Article.....	45
3) Autres méthodes de détection.....	54
a) Cartes de contrôle	54
(i) La méthode CUSUM	54
(ii) La méthode EWMA.....	57
(iii) Avantages et limites.....	60

b)	Méthodes basées sur les seuils historiques	60
(i)	Méthode de Stroup	60
(ii)	Algorithme de Farrington	62
4)	Développement d'un indicateur de suivi de la gastro-entérite par <i>Celtipharm</i>	63
a)	Présentation des indicateurs IAS®	63
b)	Indicateur de suivi de la gastro-entérite.....	63
c)	Discussion.....	64
V. Estimation de la couverture vaccinale à partir des ventes de vaccins		66
1)	Contexte et objectifs	66
2)	Vaccination contre la grippe.....	68
a)	Introduction.....	68
b)	Méthodes.....	69
c)	Population Générale.....	70
(i)	Estimation de la couverture vaccinale	70
(ii)	Délai pour la vaccination	71
d)	Population des plus de 65 ans	72
(i)	Détermination et validation de la couverture vaccinale	72
(ii)	Suivi en temps réel de la vaccination chez les plus de 65 ans	74
(iii)	Développement d'un indicateur par la société Celtipharm.....	75
e)	Discussion.....	76
f)	Conclusion	77
3)	Vaccination rotavirus	78
a)	Contexte.....	78
b)	Estimation de la couverture vaccinale	78
(i)	Méthodes.....	78
(ii)	Résultats.....	79
c)	Hospitalisations pour gastroentérites à rotavirus	80
(i)	Méthodes	80
(ii)	Résultats.....	81
d)	Corrélation entre couverture vaccinale à rotavirus et les hospitalisations.....	82
e)	Discussion.....	83
f)	Perspectives	85
VI. Autres applications du suivi des ventes de médicaments		86
1)	Evaluation des mesures de prévention en France pendant la crise nucléaire de Fukushima.	86
a)	Résumé	86
b)	Article	87
2)	Impact potentiel de la pandémie de grippe et de l'utilisation des gels hydro-alcooliques sur l'épidémie de gastro-entérite en France.....	93
a)	Résumé	93

b) Article	94
3) Utilisation des médicaments psychotropes chez les enfants et adolescents en France	101
a) Résumé	101
b) Apport des données de ventes de médicaments en pharmacies.....	101
VII. Discussion.....	102
1) Intérêt des données de ventes de médicaments par rapport aux données de surveillance existantes	102
2) Détection d'épidémies	103
3) Evaluation des mesures de santé publique	105
4) Limites	105
5) Nouvelles sources de données	106
6) Protection des données personnelles	108
VIII. Conclusion	109
Bibliographie.....	112
Liste des tableaux et Figures	120

Liste des publications et communications

Publications scientifiques liées à la thèse

M. Pivette, J.E. Mueller, P. Crépey, and A. Bar-Hen, “Surveillance of gastrointestinal disease in France using drug sales data”, *Epidemics*, vol. 8, pp. 1–8, Sep. 2014

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen, “Drug sales data analysis for surveillance of infectious diseases: a systematic literature review”, *BMC Infectious Diseases*, vol 14, no. 1, pp. 604, Nov. 2014

M. Pivette, V. Auvigne, P. Guérin, J.E. Mueller, “Development and validation of a tool to estimate influenza vaccination coverage in real time in the population aged over 65 in France”, *Soumis à la Revue d’Epidémiologie et de Santé Publique (Janvier 2015)*

P. Crépey, **M. Pivette**, A. Bar-Hen, “Quantitative Assessment of Preventive Behaviors in France during the Fukushima Nuclear Crisis”, *PLoS One*, vol. 8, no. 3, pp. 1–5, 2013

P. Crépey, **M. Pivette**, M. Desvarieux, “Potential Impact of Influenza A / H1N1 Pandemic and Hand-Gels on Acute Diarrhea Epidemic in France”, *PLoS One*, vol. 8, no. 10, pp. 1–6, 2013

Autres publications scientifiques

V. Kovess, S. Choppin, F. Gao, **M. Pivette**, M. Husky, and E. Leray. “Psychotropic Medication Use in French Children and Adolescents”, *Journal of Child and Adolescent Psychopharmacology*, January 2015

A. Boyd, S. Van de Velde, **M. Pivette**, M. Have, S. Florescu, S. O’Neill, J. Miguel Caldas-de-Almeida, G. Vilagut, J. Maria Haro, J. Alonso, V. Kovess-Masféty and the EU-WMH Investigators. “Gender differences in psychotropic use across Europe: results from a large cross-sectional, population-based study”, *European Psychiatry*, September 2015

Communications orales

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen “Drug sales data analysis for surveillance of infectious diseases: a systematic literature review”, *European Public Health Conference*, Novembre 2012, Malte.

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen “Surveillance of gastroenteritis in France using drug sales data”, *European Public Health Conference*, Novembre 2013, Bruxelles.

Communications affichées

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen “Using drug sales data for surveillance of infectious diseases: a literature review”, *Journées scientifiques du réseau doctoral de l’EHESP*, Mai 2012, Hôtel-Dieu, Paris.

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen “Surveillance of gastroenteritis in France using drug sales data”, *Journées scientifiques du réseau doctoral de l’EHESP*, Avril 2013, Ministère de la Santé, Paris.

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen “Surveillance of gastroenteritis in France using drug sales data”, *Journées scientifiques EPITER*, Avril 2013, EHESP, Rennes.

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen “Surveillance of gastroenteritis in France using drug sales data”, *Conference Epidemics*, Novembre 2013, Amsterdam.

Liste des abréviations

ACL	Association de Codification Logistique
AESPG	Association of the European Self-Medication Industry
ALD	Affection de Longue Durée
AMM	Autorisation de Mise sur le Marché
ANSM	Agence Nationale de Sécurité du Médicament et des Produits de Santé
ARS	Agence Régionale de Santé
ATC	Anatomical Therapeutic Chemical Classification
BPCO	Broncho-Pneumopathie Chronique Obstructive
CDC	Centers for Disease Control and Prevention
CIFRE	Convention Industriel de Formation par la REcherche
CIM	Classification Internationale des Maladies
CIP	Club Inter-Pharmaceutique
CNAMTS	Caisse Nationale d'Assurance Maladie des Travailleurs Salariés
CNIL	Commission Nationale de l'Informatique et des Libertés
CUSUM	Cumulative Sum Control Chart
CV	Couverture Vaccinale
DCIR	Données de consommation interrégimes
DMP	Dossier Médical Personnel
DP	Dossier Pharmaceutique
EGB	Échantillon Généraliste de Bénéficiaires
EMA	Agence Européenne des Médicaments
EphMRA	European Pharmaceutical Marketing Research Association
EWMA	Exponentially Weighted Moving Average
GEA	Gastro-entérites aiguës
HAS	Haute Autorité de Santé
HCSP	Haut Conseil de la Santé Publique
OMS	Organisation Mondiale de la Santé
OTC	Over-the-counter drugs
PMSI	Programme de Médicalisation des Systèmes d'Information
SLM	Sections Locales Mutualistes
SMR	Service Médical Rendu
SNIIRAM	Système National d'informations Inter Régions d'Assurance Maladie
SRO	Solutés de Réhydratation Orale

I. Préambule

Cette thèse a été réalisée dans le cadre d'une convention industrielle de formation par la recherche (CIFRE) (www.anrt.asso.fr) avec l'entreprise *Celtipharm* (www.celtipharm.com). Les équipes d'accueil sont le Laboratoire de Mathématiques Appliquées de Paris 5 (<http://w3.mi.parisdescartes.fr/map5/>) et le département EPIBIOSTAT de l'Ecole des Hautes Etudes en Santé Publique (www.ehesp.fr). La direction de la thèse a été réalisée par le Pr. Avner Bar-Hen (MAP 5) et la co-direction par le Dr Judith Mueller (EHESP- Département EPIBIOSTAT).

Celtipharm est une société anonyme au capital de 1 191 827 € spécialisée dans le recueil et le traitement de l'information dans le domaine de la santé (www.celtipharm.com), dont le siège social est situé 3, Allée Nicolas Leblanc à Vannes, Morbihan (56). *Celtipharm* dispose d'une antenne à Paris situé au 141, rue de Rennes. La société a été créée le 30 octobre 2000 et compte un effectif de 48 salariés. En 2013, le chiffre d'affaire de la société est de 11 166 000 €. Le Président Directeur Général de *Celtipharm* est le Dr. Patrick GUERIN. *Celtipharm* conçoit des dispositifs médico-économiques et des plans d'actions marketing-ventes pour les différents acteurs de santé. Leurs clients sont les laboratoires pharmaceutiques, des groupements d'industriels et des institutions de Santé notamment l'Agence Nationale de Sécurité des Produits de Santé (ANSM) et la Haute Autorité de Santé (HAS).

Une convention de collaboration de recherche a été signée entre *Celtipharm* et l'EHESP en 2011 pour une durée de 5 ans. L'objectif commun était de répondre à des questions de santé publique à partir de l'exploitation des données de ventes de médicaments produites par *Celtipharm* et de mener des projets de recherche en épidémiologie (notamment étudier les évolutions de consommations médicamenteuses, mener des études comparatives internationales en pharmacoépidémiologie, mettre au point des indicateurs de suivi de maladies). *Celtipharm* met à disposition de l'EHESP les données issues de son panel Xpr-SO® (échantillon de 3004 pharmacies en France Métropolitaine). La thèse est issue de cette collaboration et a pour sujet la surveillance des maladies infectieuses à partir des ventes de médicaments en France.

II. Introduction

Nous nous intéresserons dans cette thèse à déterminer l'intérêt du suivi des ventes de médicaments pour la surveillance des maladies infectieuses en France. Dans ce chapitre, nous définirons tout d'abord les objectifs de la surveillance épidémiologique et nous présenterons les avantages potentiels d'une surveillance à partir des ventes de médicaments par rapport aux autres systèmes de surveillance existants. Nous décrirons ensuite les caractéristiques du marché du médicament en France, la classification des médicaments, et la base de données de ventes de médicaments (Xpr-SO®) développée par la société *Celtipharm* qui sera utilisée dans les analyses. Nous présenterons ensuite la problématique scientifique et les axes de recherches développés dans la thèse qui seront une revue de la littérature sur la surveillance des maladies infectieuses à partir des ventes de médicaments, une étude sur la détection des épidémies de gastro-entérites en France à partir des ventes de médicaments et une étude sur l'estimation de la couverture vaccinale à partir des ventes de vaccins.

1) La surveillance épidémiologique

a) Définition

La surveillance épidémiologique a été définie par le Centers of Disease Control and Prevention (CDC) comme « un processus de collecte, de compilation et d'analyse de données, ainsi que leur diffusion à l'ensemble de ceux qui ont besoin d'être informés » (Langmuir 1963; Buehler et al. 2004). La définition implique que les données soient recueillies de façon systématique et continue, avec une diffusion rapide des résultats en vue d'une action de santé publique. La finalité d'un système de surveillance est de pouvoir informer rapidement les autorités de santé, décideurs, professionnels de santé, usagers en cas de menace afin que soient mises en place des mesures et actions pour atténuer l'impact de cette menace. La surveillance épidémiologique présente trois objectifs principaux : décrire, alerter et évaluer. Tout système de surveillance doit pouvoir caractériser un phénomène de santé dans une population en termes de temps, lieu et personne. Certains systèmes auront un objectif d'alerte, c'est-à-dire la détection de phénomènes aigus (épidémies, maladies émergentes) qui nécessitent une réponse rapide. D'autres systèmes auront pour finalité d'évaluer l'impact de programmes de santé publique et de mesures de contrôle et de prévention (Astagneau and Ancelle 2011).

b) Surveillance directe

La surveillance directe des maladies est une surveillance basée classiquement sur des diagnostics de la maladie. Elle inclut notamment la surveillance des maladies à déclaration obligatoire qui sont 31 maladies qui doivent être notifiées par les médecins et biologistes, les réseaux de médecins qui surveillent de façon volontaire un nombre défini de pathologies comme le réseau Sentinelles (Costagliola et al. 1991; Flahault et al. 2006), les réseaux de laboratoires (ex : le réseau Renago qui surveillent les infections à gonocoques), les Centres Nationaux de Référence, les registres de pathologies. Ces données permettent d'avoir des informations précises sur la maladie et son évolution, car reposant sur des diagnostics établis. Elles sont nécessaires pour confirmer les tendances de maladies ou mettre en évidence les agents pathogènes responsables. Cependant, leur réactivité peut dans certains cas être limitée, avec des délais de notifications importants. De plus, elles ne permettent de couvrir que les pathologies ciblées par la surveillance. La sensibilité peut également être limitée, notamment pour les Déclarations Obligatoires qui reposent sur le rapport de cas par les médecins. Enfin, elles nécessitent un temps et une charge de travail importante pour la collecte des données.

c) Surveillance syndromique

Pour pallier aux limites des systèmes de surveillance traditionnels pour la plupart peu adaptés à la détection précoce d'évènements sanitaires, la surveillance syndromique s'est développée à la fin des années 1990 (Fouillet et al. 2014). La surveillance syndromique est définie par le CDC comme « une approche dans laquelle les intervenants sont assistés par des procédures d'enregistrement automatiques des données, qui permettent la mise à disposition de données pour le suivi et l'analyse épidémiologique en temps réel ou proche du temps réel. Cela afin de détecter des évènements habituels ou inhabituels plus tôt qu'il n'aurait été possible de le faire sur la base des méthodes traditionnelles de surveillance » (Buehler et al. 2004). Elle s'est d'abord développée aux Etats-Unis avec comme objectif principal de pouvoir rapidement détecter des attaques bioterroristes (Buehler et al. 2003). Son utilisation s'est ensuite élargie, avec notamment la surveillance des maladies infectieuses. En France, la surveillance syndromique est apparue à partir de 2004 avec la mise en place du système SurSaud® (Surveillance Sanitaire des Urgences et des Décès), suite à la canicule de l'été 2003 (Josseran et al. 2009; Caserio-Schönemann et al. 2014). La canicule de 2003 avait mis en évidence l'absence d'outils de détection précoce d'évènements inhabituels et avait montré la nécessité

de mettre en place des systèmes permettant d'identifier en temps-réel des menaces afin de pouvoir mettre en place rapidement des mesures de contrôle et de gestion de crise. La surveillance syndromique s'est ensuite développée grâce à l'accès à des volumes importants de données reliées à la santé, notamment grâce à l'essor du numérique et le développement des méthodes de collecte et de traitement de données massives (Big data). Les données utilisées sont variées (Berger et al. 2006) et ne sont pas initialement prévues pour la surveillance sanitaire. Il peut s'agir de données en lien avec le système de santé, telles que les admissions aux urgences (Lewis et al. 2002; Josseran et al. 2010), des données provenant des ambulances (Mostashari et al. 2003), du SAMU, ou les ventes de médicaments sur ordonnance (Sugawara et al. 2012). D'autres données ne sont pas liées à une prise en charge médicale du patient. Elles sont donc moins spécifiques mais peuvent être obtenues de façon plus précoce. Il peut s'agir du taux d'absentéisme dans les entreprises ou les écoles (Kara et al. 2011), les appels à des lignes téléphoniques de conseil (Rodman et al. 1998; Rolland et al. 2006), les recherches d'information sur Internet (Pelat et al. 2009) ou les ventes de médicaments sans ordonnance (Davies and Finch 2003). La Figure 1 synthétise le type de données utilisées ainsi que leur réactivité et leur spécificité.

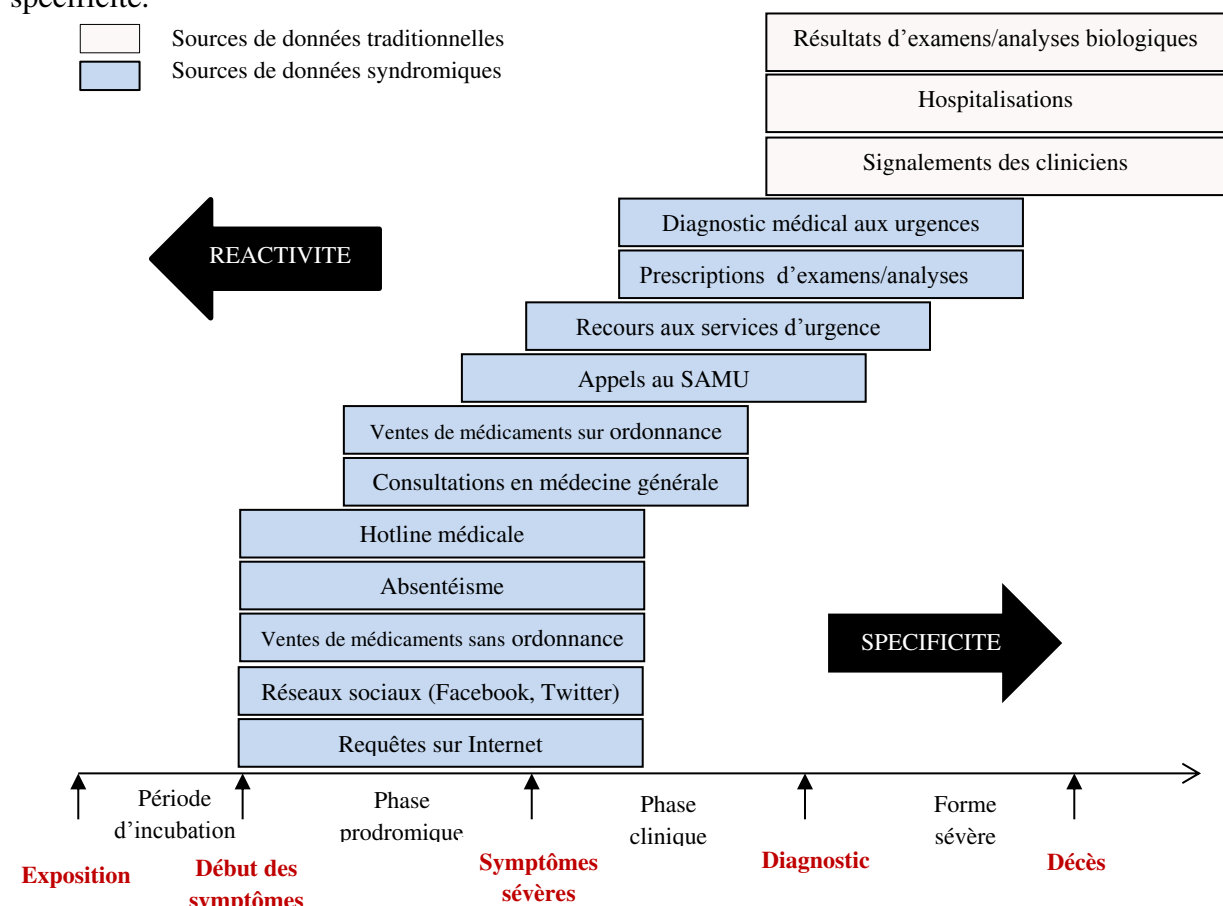


Figure 1 : Sources de données pour la surveillance épidémiologique (Invs 2013)

Les données de surveillance syndromique sont donc moins spécifiques, et s'attachent à détecter des phénomènes à partir de « proxy » de l'état de santé de la population. Elles ne sont donc pas adaptées à la surveillance de maladies peu fréquentes. Néanmoins, elles permettent d'obtenir des informations en temps réel ou proche du temps réel. Elles sont donc adaptées à la détection rapide d'évènements sanitaires et permettent la mise en place de réponses rapides en cas d'alerte. Elles se basent sur des données existantes, souvent collectées de façon automatique. Ainsi, la charge de travail et le coût est moindre par rapport aux systèmes de surveillance traditionnels. Enfin, les données sont collectées sans sélection a priori, permettant le suivi de multiples pathologies (Caserio-Schönemann et al. 2014).

d) Surveillance à partir des ventes de médicaments

Le suivi des ventes de médicaments, qui va être étudié dans la thèse, est un type de surveillance non spécifique ayant pour objectif d'évaluer l'état de santé de la population à partir de l'analyse des ventes d'une sélection de médicaments. Un des premiers exemples de l'intérêt de cette surveillance a été démontré lors d'une épidémie de *Cryptosporidium* en 1993 à Milwaukee (Canada). Les ventes d'anti-diarrhéiques hors prescription en pharmacie avaient été trois fois supérieures à la normale pendant l'épidémie, et il a été montré que cette augmentation aurait pu être mise en évidence plusieurs semaines avant que les résultats de laboratoires ne soient disponibles (Proctor et al. 1998). Cet évènement a montré l'utilité d'une surveillance non spécifique permettant de déclencher rapidement une alerte et d'informer les autorités. La surveillance à partir des ventes de médicaments s'est ensuite développée à partir des années 2000 avec la mise en place de systèmes automatisés de collecte de données de ventes (Chen et al. 2010). On peut citer le Real-Time Outbreak and Disease Surveillance system (Kirian and Weintraub 2010) qui est un système de surveillance développé par l'Université de Pittsburg en 1999 et dont une des composantes est le suivi quotidien des ventes de 18 groupes de médicaments « Over-the-Counter (OTC) » issues de 29000 pharmacies dans dix états américains. Le système de surveillance de la ville de New York a lui mis en place en 2002 la collecte quotidienne des ventes de médicaments indiqués pour des syndromes grippaux et gastro-intestinaux issues de 248 pharmacies à New York (Das et al. 2005). Le système de surveillance ESSENCE, développé en 2002, collecte les ventes issues de 300 pharmacies dans trois états américains (Maryland, Washington, Virginie) (Lombardo et al. 2003; Magruder et al. 2004). Dans tous ces systèmes de collection automatique de

données, le délai de transmission des données est d'un jour et l'analyse quotidienne des données permet de mener des investigations en cas de détection d'anomalies.

Les données de ventes de médicaments apparaissent comme un outil intéressant pour la surveillance épidémiologique en France. La rapidité d'obtention des données peut permettre l'évaluation rapide de l'état de santé de la population et la détection d'évènements inhabituels. Par exemple, les données de ventes de médicaments issues du panel Xpr-SO® de l'entreprise *Celtipharm* (3004 pharmacies en France Métropolitaine) sont collectées automatiquement et disponibles avec un délai d'un jour (*Celtipharm*). De plus, les données de ventes ne sont pas collectées spécifiquement pour une pathologie mais contiennent 100% des ventes réalisées en pharmacie, elles semblent donc être un outil pertinent pour détecter l'apparition de maladies nouvelles ou émergentes. Le volume de données recueillies est également important et permet de couvrir une large portion de la population (en moyenne 2925 habitants par pharmacie en France). Enfin, les données de ventes apportent une information directement liée à l'état de santé du patient et sont donc plus spécifiques que des sources de données de surveillance indirecte telles que les requêtes sur Internet ou les messages sur les réseaux sociaux (Aslam et al. 2014).

Les données de ventes que nous avons analysées dans cette thèse permettent de différencier les médicaments prescrits et les médicaments achetés sans ordonnance. Les ventes réalisées sans ordonnance peuvent permettre d'identifier les patients présentant des symptômes légers qui achètent des médicaments sur le conseil du pharmacien et qui ne sont donc pas enregistrés dans des systèmes basés sur les consultations médicales, les recours aux urgences ou les hospitalisations. L'accès à des données de ventes de médicaments achetés sans ordonnance pourrait également présenter un intérêt pour la détection et l'alerte précoce. Ils peuvent permettre d'identifier précocement des évènements de santé ; les patients pouvant acheter les médicaments non-prescrits lors des premiers symptômes et avant une consultation médicale. Concernant les médicaments prescrits qui sont liés au système de soins car achetés après une consultation médicale, ils peuvent présenter un intérêt pour la surveillance des maladies non surveillées ou pour lesquelles les données cliniques sont difficiles à obtenir ou coûteuses.

Il existe actuellement en France des données sur la consommation de médicaments obtenues à partir de bases administratives et médico-administratives. Le SNIIRAM (Assurance Maladie 2015) est un entrepôt de données anonymes regroupant les informations issues des

remboursements effectués par l'ensemble des régimes d'assurance maladie pour les soins du secteur libéral. Trois ensembles de restitution sont disponibles : des données agrégées appelées datamarts, un échantillon général des bénéficiaires (EGB) au 1/97^{ème} de la population assurée, une base exhaustive de données individuelles des bénéficiaires (DCIR) pour réaliser des études sur la consommation de soins. Les données du SNIIRAM ont par exemple été utilisées pour décrire l'épidémiologie de certaines maladies telles que le diabète (Kusnik-Joinville et al. 2008), la maladie d'Alzheimer (Tuppin et al. 2012b) ou de Parkinson (Blin et al. 2014). Elles permettent de suivre dans le temps la consommation médicamenteuse des patients grâce à un numéro d'identifiant du patient et de pouvoir identifier de façon spécifique les patients présentant des pathologies. Cependant, le délai d'obtention peut être long (parfois plusieurs années) et la base de données n'est pas adaptée à la surveillance en temps réel. Elle est donc peu adaptée aux maladies infectieuses. Les études existantes portent en effet uniquement sur des maladies chroniques (Martin-Latry 2010). De plus, elle donne accès uniquement aux produits remboursés par l'Assurance Maladie, et ne permet pas de connaître les achats de produits non remboursables. Le suivi des ventes de médicaments apparaît donc comme un outil plus réactif et plus complet, l'analyse des ventes de médicaments permettant le suivi des médicaments achetés sans ordonnance.

2) La base de données *Celtipharm* et le médicament en France

a) Le panel Xpr-SO®

Les données de ventes analysées dans nos travaux sont issues la base de données Xpr-SO® (*Celtipharm*) développée par l'entreprise *Celtipharm*. Il s'agit d'un échantillon aléatoire stratifié de 3004 pharmacies en France Métropolitaine, soit 13% du nombre total de pharmacies (21614 pharmacies). Les données de ventes achats et stocks de 100% des produits (médicaments et parapharmacie) des pharmacies de l'échantillon sont récupérées de façon automatique plusieurs fois par jour et accessible le lendemain (J+1). Les données sont disponibles depuis 2008. *Celtipharm* met à jour de façon continue une base de données sur les caractéristiques de l'ensemble des officines de France Métropolitaine. Une stratification de l'ensemble des pharmacies a été réalisée à partir :

- du chiffre d'affaire (6 niveaux pour le chiffre d'affaire global, 4 niveaux par types de ventes : médicaments à prescription obligatoire, médicaments à prescription facultative, autres produits)
- de la localisation (5 aires géographiques)
- de la zone de chalandise (5 niveaux, de rural à fortement urbanisé)

Chaque strate a un minimum de 30 pharmacies ou est fusionnée avec la strate la plus proche. Le taux d'échantillonnage par strate a été calculé avec l'algorithme d'allocation optimale de Neyman (Kish 1965). Les données recueillies à partir de l'échantillon des 3004 pharmacies sont ensuite extrapolées à la France entière à partir des strates. La méthode d'extrapolation à partir de l'échantillon a été conçue en collaboration avec le département statistique et Informatique Décisionnelle de l'Université de Bretagne Sud. Les extrapolations ont été validées à partir de l'analyse des ventes de produits des industriels aux officines de ville.

b) Utilisation de la base de données pour des études épidémiologiques

Dans le cadre du contrat de collaboration de la thèse CIFRE, *Celtipharm* met à disposition les données de son panel Xpr-SO®. Les données disponibles pour des analyses épidémiologiques et qui seront utilisées dans nos travaux sont le nombre de boîtes de médicaments et produits unitaires vendus en officines et définis par leur Code Identifiant de Présentation (CIP) s'il s'agit d'un médicament ou de leur code ACL (Association de Codification Logistique) pour les autres produits de santé. Les données sont disponibles de façon journalière. Elles sont extrapolées au niveau national et régional (France Métropolitaine, hors Corse). Le statut du médicament en termes de prescription est enregistré, on peut ainsi différencier les produits achetés avec une ordonnance et ceux achetés sans ordonnance et uniquement avec le conseil d'un pharmacien. L'année de naissance du patient est renseignée pour les médicaments prescrits ou achetés en même temps que des médicaments prescrits. En effet, l'enregistrement de la carte vitale du patient par la pharmacie lors de l'achat permet de connaître l'année de naissance. Les données sont disponibles depuis janvier 2008.

Cependant, il faut noter que ces données ne permettent pas de suivre la consommation dans le temps des patients car il n'existe pas d'identifiant unique qui permettrait de les identifier lors de chaque passage en pharmacies. Il n'y a également pas d'information sur les antécédents médicaux des patients comme leur statut ALD. Les travaux réalisés vont porter sur un nombre de boîtes de médicaments vendues, et il ne sera donc pas possible de déterminer une incidence

mais uniquement un nombre de boîtes de médicaments vendus par personne. Un nombre de patients pourra néanmoins être déterminé dans le cas où un seul médicament est indiqué par personne, comme dans l'étude sur la vaccination grippale où une seule dose de vaccin est nécessaire par personne.

c) Les pharmacies en France

Le nombre d'officines en France Métropolitaine est de 21614 pharmacies au 1^{er} février 2015 (Ordre National des pharmaciens 2015a). La réglementation sur l'installation des officines a pour objectif d'avoir une répartition homogène des pharmacies sur le territoire selon le nombre d'habitants et d'assurer à la population un service de proximité. L'ouverture d'une officine n'est donc pas libre et est encadrée par la loi en fonction de la population communale et soumise à la délivrance d'une licence par l'agence régionale de santé (ARS). L'installation d'une première officine est autorisée dans les communes de plus de 2 500 habitants, l'installation d'officines supplémentaires est autorisée par tranche de 4 500 habitants. En janvier 2015, le nombre moyen d'habitants par officine était de 2925 habitants par officine en France (Ordre National des pharmaciens 2015b). Cependant la densité par habitant est légèrement variable d'une région à l'autre, elle est plus faible dans les régions du nord de la France (1 pharmacie/4043 habitants en Alsace). Dans les régions du sud de la France où la population est plus âgée et les régions où l'attractivité touristique est importante, la densité pharmaceutique est plus forte (1 pharmacie/ 2591 habitants en Provence-Alpes-Côte d'Azur).

d) Le marché des produits officinaux en France

Les études développées dans cette thèse portant sur les ventes de médicaments en France, nous décrivons dans cette partie l'organisation spécifique du marché français du médicament, les différents types de produits vendus en pharmacies de ville ainsi que leurs caractéristiques en termes de prescription et de remboursement. Le champ des produits vendus en pharmacie se divise en deux catégories : les médicaments et les autres produits de santé (dispositifs médicaux, parapharmacie vétérinaire, diététique spécialisée, cosmétique, parapharmacie classique).

(i) Le médicament : définition et classification

Le médicament est défini d'après l'article L.5111-1 du Code de la santé publique comme « toute substance ou composition présentée comme possédant des propriétés curatives ou

préventives à l'égard des maladies humaines ou animales, ainsi que tout produit pouvant être administré à l'homme ou l'animal, en vue d'établir un diagnostic médical ou de restaurer, corriger ou modifier leurs fonctions organiques ». Un médicament est soit une spécialité pharmaceutique, soit une préparation magistrale ou officinale (médicaments préparés en pharmacie). La spécialité pharmaceutique est le nom donné au médicament fabriqué industriellement, selon un mode de fabrication défini et contrôlé. Toute spécialité pharmaceutique doit faire l'objet avant sa commercialisation d'une Autorisation de Mise sur le Marché délivrée par l'Agence National de Sécurité du Médicament (ANSM) ou l'Agence Européenne des Médicaments (EMA). La délivrance d'un médicament ne peut s'effectuer en ville que dans l'une des 21614 pharmacies et sous la responsabilité d'un pharmacien. En 2013, 9900 spécialités différentes étaient commercialisées dans les pharmacies de ville en France (ANSM 2014). Chaque présentation d'une spécialité pharmaceutique est identifiée par un code numérique CIP de 7 chiffres.

En France, la prescription médicale est obligatoire pour toute spécialité inscrite sur une liste (article L.5132-6) : liste I (médicaments dangereux), liste II (médicaments dangereux moins toxiques que ceux de la liste I), et liste des stupéfiants. Toutes les spécialités ne présentant pas les critères d'inscription sur une liste mentionnée sont des médicaments à prescription facultative.

Pour être pris en charge par l'Assurance Maladie, un médicament doit être inscrit sur la liste des spécialités pharmaceutiques remboursables aux assurés sociaux. La mission d'examen des médicaments est réalisée par la commission de la Transparence de la HAS (Haute Autorité de Santé). Les médicaments sont classés en plusieurs catégories en fonction de leur service médical rendu (SMR) (Médicaments irremplaçables, SMR majeur, SMR modéré, SMR faible). De ce classement dépend le remboursement d'un médicament (respectivement 100%, 65%, 30%, 15%). Les médicaments dont le SMR est insuffisant au regard des autres médicaments ou thérapies disponibles ne sont donc pas remboursables.

L'organisation du marché français du médicament peut être décrite en fonction du statut de remboursement et de prescription (Figure 2). Les médicaments d'« automédication » correspondent aux médicaments achetés sans prescription médicale.

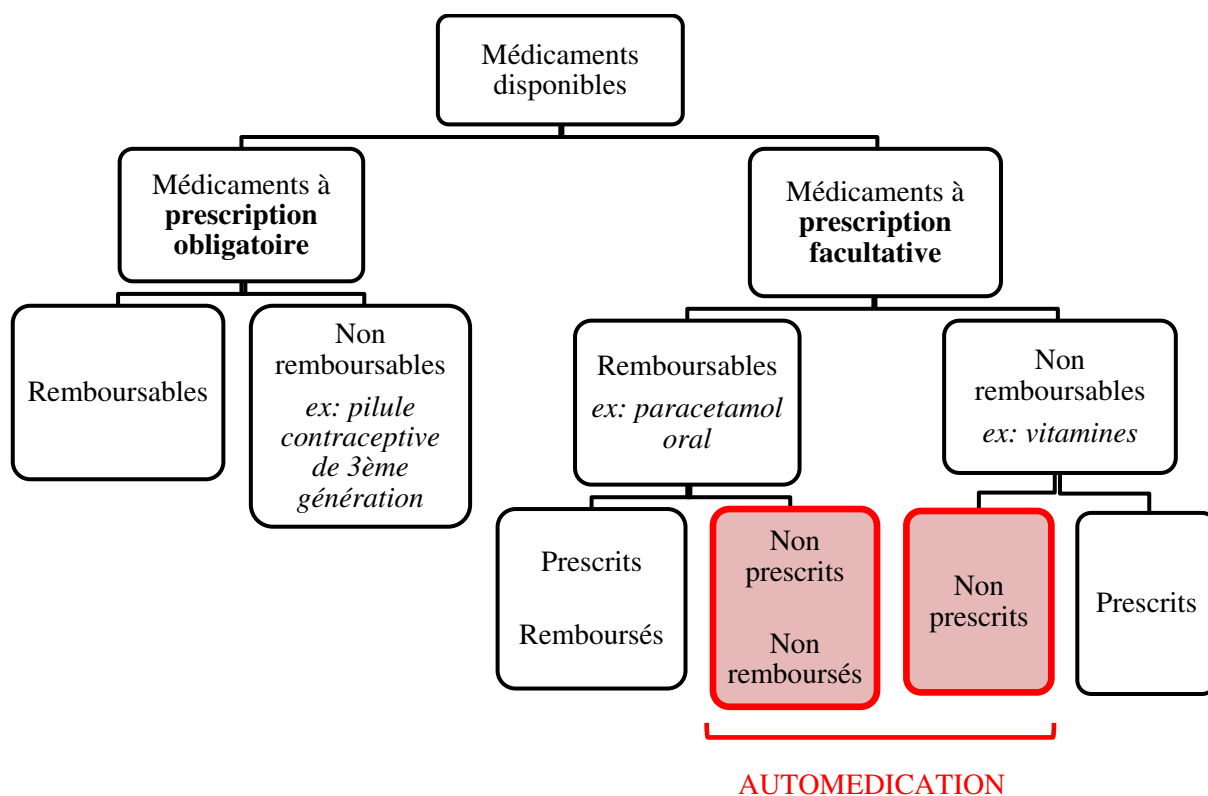


Figure 2 : Organisation du marché français du médicament.

(ii) Les chiffres du marché officinal du médicament

La décomposition du marché officinal en quantités de vente et en chiffre d'affaire en 2013 est présentée en Figure 3 (ANSM 2014). Il s'agit des ventes de médicaments effectuées auprès des officines et déclarées par les entreprises pharmaceutiques. En 2013, le chiffre d'affaire réalisé par la vente de spécialités pharmaceutiques était de 20,6 milliards d'euros et le nombre de boîtes vendues de 3,080 millions. Les spécialités à prescription obligatoire (SPO) représentent près de 54% des spécialités vendues (1,663 millions de boîtes) et 84% du chiffre d'affaire (17,3 milliard d'euros). La plupart sont remboursables (52,3% des ventes) car les pathologies graves prises en charge par les régimes d'assurance maladie sont traitées par des médicaments nécessitant une prescription médicale. Les spécialités non remboursables soumises à prescription (1,7% des ventes) sont essentiellement constituées par les contraceptifs oraux, les traitements des troubles de l'érection et quelques vaccins. Les spécialités à prescription facultatives (SPF) représentent 46% des spécialités vendues (1,416 millions de boîtes) et 16% du chiffre d'affaire (3,296 milliard d'euros). Près d'un tiers des spécialités à prescription facultative ne sont pas remboursables (997 000 boîtes).

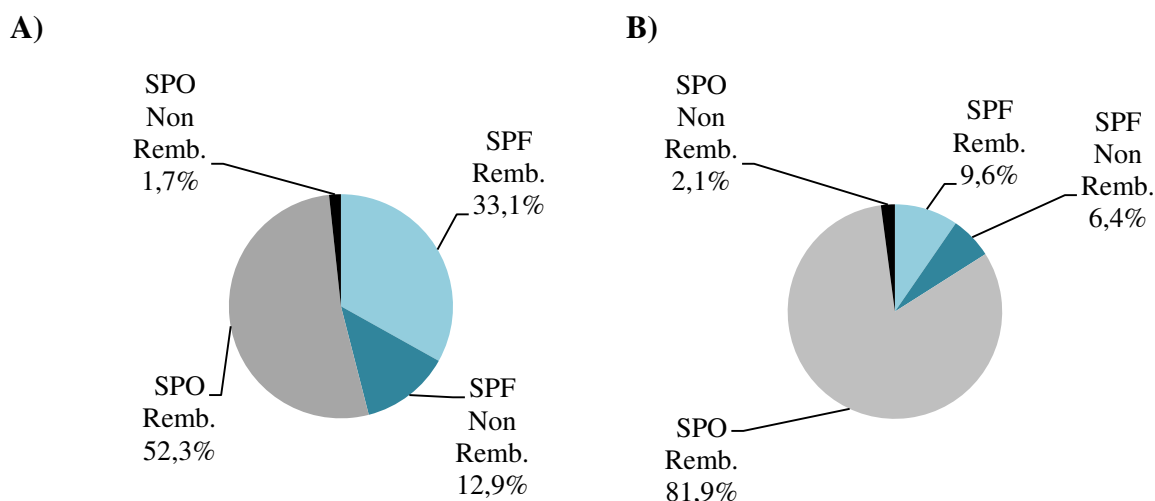


Figure 3: Décomposition 2013 du marché officinal selon le statut des spécialités (ANSM 2014) A) En quantité B) En valeur

Les médicaments d'automédication sont les médicaments achetés sans ordonnance et uniquement conseillés par un pharmacien. Il s'agit de spécialités à prescription facultative qui sont non prescrites et donc non remboursées. Dans les travaux développés dans nos études, nous distinguerons deux types de médicaments : les médicaments achetés avec une ordonnance et les médicaments achetés sans ordonnance qui correspondent donc à l'automédication. Ces médicaments d'automédication représentaient en 2013 en volume 15% des ventes totales de médicaments (465 000 boîtes vendues) effectuées par les officines et 8% de leur chiffre d'affaire (2,123 milliards d'euros) (AFIPA and Celtipharm 2014).

(iii) Autres produits de santé

Tous les produits de santé autres que les médicaments (dispositifs médicaux, parapharmacie vétérinaire, diététique spécialisée, cosmétique, parapharmacie classique) sont à prescription facultative et sont accessibles librement. Ils peuvent être également vendus dans d'autres structures que les officines. Le code ACL est un numéro d'identification de ces produits de santé.

e) La classification des médicaments

Deux types de classification des médicaments sont utilisés : la classification ATC (Anatomical Therapeutic Chemical) développée par l'OMS et la classification AT (Anatomical Therapeutic) développée par l'European Pharmaceutical Marketing Research Association (EphMRA). La classification ATC se base sur les substances actives, et en

fonction de leur action pharmaceutique et thérapeutique. La classification EphMRA se base sur les produits, en fonction de leur indication et de leur usage. La classification EphMRA comporte trois niveaux : anatomique, pharmaco-thérapeutique et chimique. Par exemple, le médicament ERYPHAR® 1000 mg est un antibactérien qui présente la classification ci-dessous :

- classe anatomique (niveau 1) : J Anti-infectieux généraux à usage systémique
- classe pharmaco-thérapeutique (niveau 2) : J01 Antibactériens à usage systémique
- classe chimique (niveau 3) : J01F Macrolides et apparentés

La classification EphMRA est la classification utilisée par les industries pharmaceutiques. Il s'agit de la classification qui sera utilisée pour sélectionner les médicaments dans notre étude sur la surveillance de la gastro-entérite en France.

3) Problématique scientifique

Comme nous l'avons montré précédemment, les ventes de médicaments présentent des avantages potentiels pour la surveillance épidémiologique grâce au volume important de données obtenues, l'exhaustivité, la rapidité d'obtention et le faible coût de celles-ci.

La possibilité d'obtenir des données en quasi temps-réel en font des données particulièrement intéressantes pour la détection d'épidémies et l'alerte sanitaire. Ces données semblent ainsi adaptées à la surveillance des maladies infectieuses pour lesquelles il est important d'avoir une détection rapide afin de pouvoir mettre en place des mesures de prévention et de contrôle. De plus, l'absence de suivi dans le temps des ventes en pharmacies limite l'utilisation de ces données pour la surveillance des maladies chroniques, pour lesquelles il est nécessaire de pouvoir identifier les patients pour estimer une prévalence. Nous avons donc limité nos analyses aux maladies infectieuses.

L'objectif de ce projet de thèse était de déterminer l'intérêt du suivi des ventes de médicaments pour la surveillance des maladies infectieuses en France. Le projet de thèse a été structuré selon trois axes de recherche qui vont être développés:

Nous avons vu qu'il existait quelques systèmes de surveillance qui utilisaient le suivi des ventes de médicaments en routine, cependant aucun état de l'art n'avait été réalisé sur l'intérêt du suivi des ventes de médicaments pour la surveillance des maladies infectieuses. Une première étape du projet a donc consisté à réaliser une revue systématique de la littérature sur la surveillance des maladies infectieuses à partir des ventes de médicaments, afin de

déterminer la validité, l'utilité ainsi que les limites de cette approche. Cette étude a permis de décrire le potentiel de ces données pour la détection d'épidémies en temps réel et d'identifier les pathologies adaptées à une surveillance par les médicaments. Cette première étape a également permis d'analyser les différentes méthodes utilisées pour choisir les médicaments traceurs et les différentes méthodes de détection d'épidémies. Les résultats ont été présentés en fonction du type de médicaments, prescrits et non-prescrits, et nous avons ainsi pu analyser l'intérêt des médicaments non prescrits pour la détection précoce des épidémies.

Suite à cette revue de la littérature, nous avons réalisé une étude sur la détection des épidémies saisonnières de gastro-entérites en France à partir des ventes d'un groupe de médicaments. L'étude a porté sur la gastro-entérite car il s'agit d'un problème de santé publique (en moyenne 1 700 000 cas par saison hivernale en 2008-2013) pour lequel une détection rapide et valide est nécessaire. Il est également possible d'identifier des médicaments indiqués en cas de gastro-entérite et suffisamment spécifiques (antidiarrhéiques, antiémétiques, absorbants, SRO etc.) qui permettront le suivi de la maladie. De plus, une partie des médicaments utilisés dans le traitement de la gastro-entérite sont à prescription facultative et nous pourrions donc analyser séparément les médicaments non-prescrits afin de déterminer leur performance pour la détection précoce des épidémies. La gastro-entérite est également une maladie pertinente à suivre car il existe en France un réseau de médecins généralistes qui rapportent le nombre de patients vus en consultation pour gastro-entérite (French GPs Sentinelles Network). Ces données de référence pourront donc être utilisées pour valider les signaux obtenus à partir des ventes de médicaments. L'objectif de notre étude était de déterminer la validité d'une surveillance de la gastro-entérite en France par les ventes de médicaments en les comparant à des données de référence, et d'identifier le potentiel de ces données pour la détection précoce d'épidémies.

Une troisième étude a porté sur l'estimation de la couverture vaccinale (CV) et le suivi en temps réel de l'activité vaccinale en France à partir des ventes de vaccins. Contrairement aux deux études précédentes, ce ne sont pas ici les ventes de médicaments utilisés pour le traitement de maladies infectieuses qui ont été analysées mais les ventes de vaccins utilisés en prévention de maladies infectieuses. Il est important d'avoir des données de couverture vaccinale dans la population en quasi temps-réel afin d'évaluer si les programmes de vaccination sont correctement appliqués et pouvoir ajuster les messages de communication. Les données de ventes de vaccins peuvent potentiellement présenter un gain important en terme de réactivité et de représentativité par rapport aux sources de données existantes

(certificats de santé, enquêtes en population etc.). Notre étude a ciblé deux vaccins : le vaccin antigrippal et le vaccin rotavirus. Les données de couverture vaccinale grippe sont actuellement communiquées après la campagne de vaccination et concernent uniquement les vaccins remboursés aux assurés du Régime Général. Notre objectif était d'estimer la couverture vaccinale de l'ensemble de la population et de fournir une information en temps réel. Le vaccin rotavirus est indiqué chez les nourrissons de moins de 6 mois. Il n'est pas remboursé et il n'existe pas actuellement d'estimation de la couverture vaccinale en France. L'objectif de notre étude était d'estimer la couverture vaccinale rotavirus en France depuis 2008 et d'analyser l'association avec le nombre d'hospitalisations pour gastro-entérite chez les enfants de moins de 2 ans.

III. Revue de la littérature sur la surveillance des maladies infectieuses à partir des ventes de médicaments

Cette revue de la littérature a fait l'objet d'une publication scientifique en novembre 2014 dans la revue *BMC Infectious Diseases* :

M. Pivette, J.E. Mueller, P. Crépey, A. Bar-Hen, "Drug sales data analysis for surveillance of infectious diseases: a systematic literature review". *BMC Infectious Diseases*, vol 14, no. 1, pp. 604, 2014

1) Résumé

Nous avons réalisé une revue systématique de la littérature sur la détection des maladies infectieuses à partir de l'analyse des données de ventes de médicaments. L'objectif était de déterminer la validité et l'utilité de l'analyse des ventes de médicaments ainsi que les limites et avantages de cette approche. Nous avons limité la revue de la littérature aux maladies infectieuses pour lesquelles une détection précoce et valide est nécessaire afin de pouvoir mettre en place rapidement des mesures de contrôle.

Une recherche de publications 'peer-reviewed' a été réalisée sur les bases de données Pubmed, Scopus, Embase, Cochrane Library, African Index Medicus, et Lilacs à partir des mots-clés ("surveillance" OR outbreak detection OR warning system) AND (over-the-counter OR "prescription drugs" OR pharmacy OR (pharmaceutical OR drug OR medication) sales). La sélection d'articles a été réalisée de 1975 à juin 2012, sans limitation de lieu, en langue anglaise ou française. Nous avons sélectionné des articles qui décrivaient, analysaient une surveillance de maladies infectieuses à partir des ventes de médicaments. La méthodologie PRISMA a été suivie. Les articles sélectionnés ont été analysés en termes d'objectifs, de maladies surveillées, de sources de données, de méthodologies et de performance pour la détection d'épidémies en temps réel. Pour décrire les méthodes et les résultats, nous avons séparé les articles en trois groupes : les études rétrospectives, les études visant à sélectionner des médicaments traceurs et les études prédictives. Les indicateurs utilisés pour comparer les données de ventes et les données de référence étaient : la corrélation entre les séries temporelles (force et délai) et la performance pour la détection d'épidémies (sensibilité, spécificité et délai de détection).

3266 articles ont été identifiés à partir de la recherche par mots-clés. 110 ont ensuite été sélectionnés à partir de la lecture du titre et du résumé. La lecture des 110 articles a permis d'identifier 27 articles correspondant aux critères d'inclusion et qui ont été analysés dans cette revue de la littérature. La plupart des études s'intéressaient aux syndromes grippaux (17 études) et la gastro-entérite (11 études). Les données de médicaments étaient collectées spécifiquement pour l'étude à partir d'un échantillon de pharmacie (11 études) ou pouvaient être collectées en routine par des laboratoires publics, des agences de santé ou des entreprises privées spécialisées dans la collecte de données de santé (16 études). Dix-neuf articles étaient des études rétrospectives descriptives qui comparaient les données de ventes avec des données cliniques de référence de la maladie. Dix-sept ont montré des corrélations significatives avec les données de ventes et les données de référence. Quatre études ont montré que les ventes de médicaments disponibles sans ordonnance (OTC) pouvaient donner un signal précoce d'épidémie. Cinq études ont développé et validé des méthodes de sélection des médicaments traceurs pour le suivi de maladies spécifiques. Dans trois autres études, des modèles de prédiction ont été développés pour prédire le nombre de cas à partir des ventes de médicaments.

Les résultats montrent que le suivi des ventes de médicaments est un outil efficace pour la surveillance des maladies grippales et gastro-intestinales. Les médicaments prescrits étaient très corrélés aux données cliniques, sans délai. De fortes corrélations ont été également observées entre les ventes de médicaments OTC et les données cliniques. Quelques études ont montré que les médicaments OTC permettaient une détection précoce des épidémies. Les patients achètent des médicaments sans ordonnance lors des premiers symptômes et avant une consultation médicale, ce qui permet une détection précoce des épidémies. Les limites à l'utilisation des données de ventes sont le fait que l'achat de médicaments peut être influencé par des promotions, la période de l'année, ou des annonces médiatiques. Les comportements d'achats varient également selon des facteurs sociodémographiques. De nombreux systèmes de collecte automatique de données de ventes ont été développés depuis le début des années 2000, ils illustrent la nécessité de systèmes de surveillance permettant une détection rapide d'évènements de santé. Les études portaient essentiellement sur les maladies grippales et gastro-intestinales car il existe des données cliniques de référence, seules deux études ont porté sur la surveillance de la coqueluche et de la syphilis.

Cette revue de la littérature suggère que le suivi des ventes de médicaments est une méthode prometteuse pour la surveillance et la détection des maladies infectieuses. Les données de

ventes de médicaments non-prescrits ont le potentiel pour émettre une alerte plus précocement que la plupart des systèmes de surveillance. La revue s'est limitée à la dynamique temporelle des épidémies et des analyses supplémentaires seraient nécessaires pour déterminer leur efficacité pour analyser la propagation géographique. Des bases de données sur les médicaments avec une transmission des données en quasi temps réel sont disponibles dans plusieurs pays, et des études pourraient être développées sur d'autres maladies infectieuses.

2) Article

RESEARCH ARTICLE

Open Access

Drug sales data analysis for outbreak detection of infectious diseases: a systematic literature review

Mathilde Pivette^{1,2,3*}, Judith E Mueller^{1,4}, Pascal Crépey^{1,5} and Avner Bar-Hen²

Abstract

Background: This systematic literature review aimed to summarize evidence for the added value of drug sales data analysis for the surveillance of infectious diseases.

Methods: A search for relevant publications was conducted in Pubmed, Embase, Scopus, Cochrane Library, African Index Medicus and Lilacs databases. Retrieved studies were evaluated in terms of objectives, diseases studied, data sources, methodologies and performance for real-time surveillance. Most studies compared drug sales data to reference surveillance data using correlation measurements or indicators of outbreak detection performance (sensitivity, specificity, timeliness of the detection).

Results: We screened 3266 articles and included 27 in the review. Most studies focused on acute respiratory and gastroenteritis infections. Nineteen studies retrospectively compared drug sales data to reference clinical data, and significant correlations were observed in 17 of them. Four studies found that over-the-counter drug sales preceded clinical data in terms of incidence increase. Five studies developed and evaluated statistical algorithms for selecting drug groups to monitor specific diseases. Another three studies developed models to predict incidence increase from drug sales.

Conclusions: Drug sales data analyses appear to be a useful tool for surveillance of gastrointestinal and respiratory disease, and OTC drugs have the potential for early outbreak detection. Their utility remains to be investigated for other diseases, in particular those poorly surveyed.

Keywords: Disease outbreaks, Syndromic surveillance, Drug utilization, Nonprescription drugs, Communicable diseases, Epidemiology

Background

Since the mid-1990s and the raise of concerns about bio-terrorism and emerging diseases, non-diagnosis-based data have increasingly been used for routine disease surveillance and outbreak detection [1]. The CDC defined “syndromic surveillance” as an investigational approach where health department staff, assisted by automated data acquisition and generation of statistical alerts, monitor disease indicators in real-time or near real-time to detect outbreaks of disease earlier than would otherwise be possible with traditional public health methods [2].

In such efforts, different registries have served as data sources for public health surveillance [1,3], including data on absenteeism at work or school [4], calls to health

helplines [5,6], emergency department consultations [7,8], ambulance dispatching [9], or drug sales. Although unspecific, such data sources can have the advantage over diagnosis-based surveillance of providing information within short delays since the event and in readily available electronic form for relatively low-cost, while capturing large parts of the population.

Drug sales data analysis may overcome the limitation of poor specificity when groups of drugs are exclusively used for the disease or disease syndrome of interest. Furthermore, drug sales data may earlier capture changing population health status, as over-the-counter (OTC) sales and a dense network of pharmacies in most developed countries make drugs easily accessible to patients at the earliest appearance of their symptoms.

Despite this potential interest, no state of the art of drug sales based surveillance is available to date. The

* Correspondence: mathilde.pivette@ehesp.fr

¹EHESP French School of Public Health, Sorbonne Paris Cité, Rennes, France

²Université Paris Descartes, MAP5, Paris, France

Full list of author information is available at the end of the article

present systematic literature review therefore summarized the evidence for an added value of drug sales data for infectious disease surveillance. We limited the scope of the review to infectious diseases, as they represent a public health problem for which early and valid signal detection is of particular concern, in light of potentially rapid emergence and opportunity for control interventions.

Methods

We conducted a literature search from 1975 up to June 2012 to identify relevant peer-reviewed articles regarding surveillance of infectious diseases based on drug sales data. PRISMA guidelines were followed in the reporting of the review [10]. Published articles were searched for on electronic databases (Pubmed, Embase, Scopus, LILACS, African Index Medicus, Cochrane Library), using combinations of the following key words: (“surveillance” OR outbreak detection OR warning system) AND (over-the-counter OR “prescription drugs” OR pharmacy OR (pharmaceutical OR drug OR medication) sales). The search was limited to articles in English or French. There were no limitations on study settings.

To be included in the review, articles had to describe, test, or review an infectious disease surveillance based on drug sales data; and be original research that presented

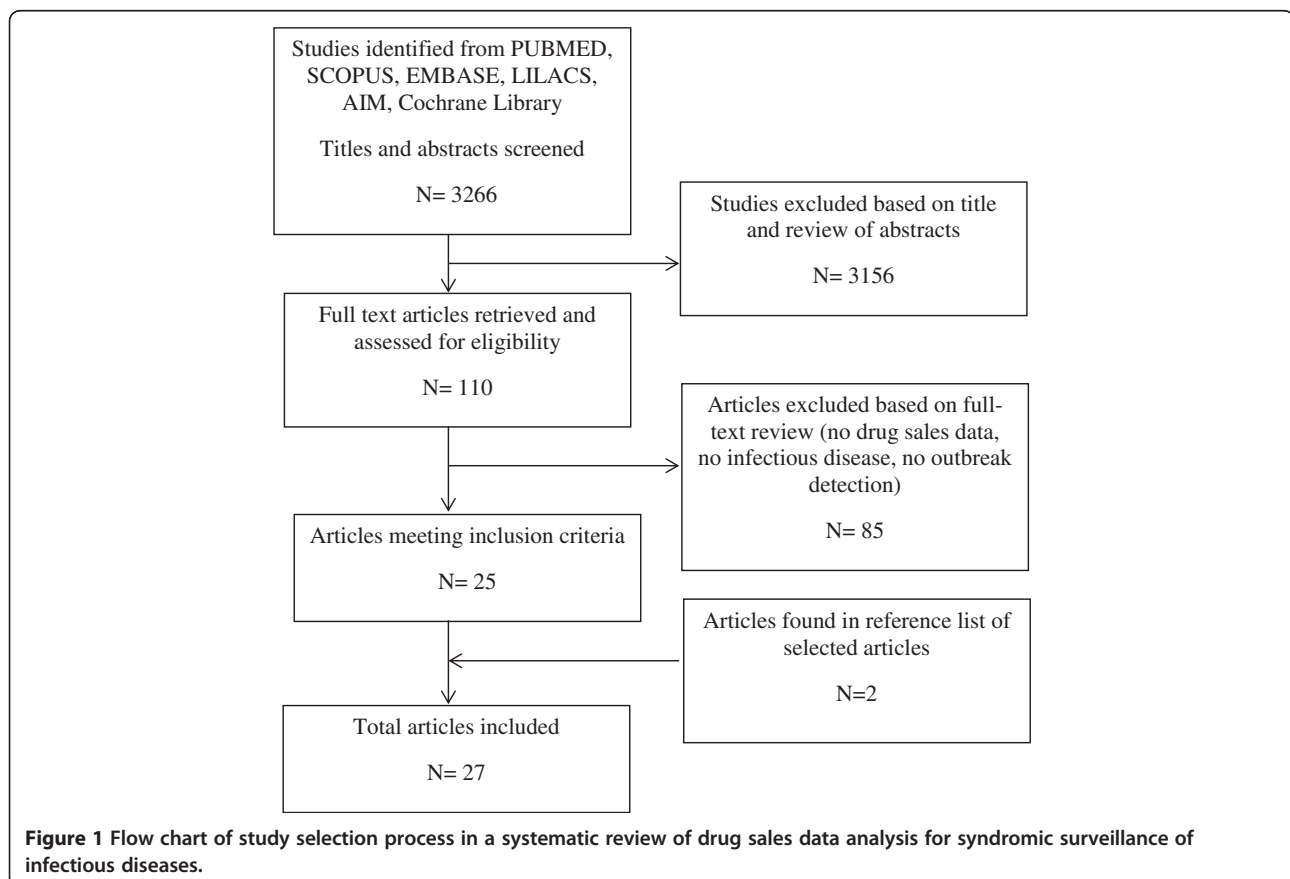
new data and results. We excluded studies that monitored chronic diseases, as well as prevalence studies whose purpose was not epidemic detection.

One reviewer screened and evaluated the titles and abstracts. Articles were widely included in a first stage. The full-text review and the final selection of the articles were made by two reviewers.

We reviewed and described the articles in terms of objectives, diseases studied, data sources, methodologies, and performance for real time surveillance. To describe methods and results, we separated the articles into three groups based on their main objective: descriptive retrospective studies, drug selection studies, and prediction studies. Outcomes selected to compare drug sales data to reference surveillance data of the corresponding disease were correlation measurements (strength and timeliness of the correlation) and indicators of outbreak detection performance (sensitivity, i.e. ability to identify true outbreaks; specificity, i.e. ability to identify true negative and timeliness of the detection).

Results

We screened a total of 3266 articles, of which 27 were included in the final review. The search and selection process is presented in Figure 1.



Objectives of the studies

Three types of studies were defined: retrospective descriptive studies, drug selection studies and prediction studies.

Nineteen of the 27 studies were descriptive retrospective studies assessing the strength of the correlation between drug sales and reference surveillance data of the corresponding disease or evaluating outbreak-detection performance [11-30]. Five studies used statistical algorithms to select groups of drugs that were closely associated with clinical surveillance data of a given disease and that would be most appropriate for future drug-sales-based surveillance [21,31-34]. In a third group of three studies, the authors developed and evaluated statistical models to predict clinical surveillance data based on drug sales data [35-37]. Table 1 summarizes the studies in terms of their general characteristics.

Diseases studied

Most of the studies focused on respiratory illnesses (17 studies) [12,14,15,18,20-22,27-32,34-37] or gastrointestinal illnesses (11 studies) [14,16-19,23-26,32,33]. Only two other studies evaluated surveillance of pertussis [11] and syphilis [13].

Populations surveyed

Most of the studies were set in the United States ($n = 16$ studies, 59%), followed by Canada ($n = 4$), France ($n = 3$), Japan ($n = 3$), the Netherlands ($n = 1$) and England ($n = 1$). Only one study was conducted in more than one country [25].

Drug sales data sources

In most retrospective studies, data were collected specifically for the purpose of the study from a sample of pharmacies [16,24-26] or from retailers [13,15,17,29]. For example, in a Canadian study [17], electronic data were provided by one major retailer for all of their 19 pharmacies in the study area.

Automatically compiled data sources were used in all the drug selection and prediction studies and in some retrospective studies. Drug sales data were routinely collected in samples of a city's or country's pharmacies. Such routine data collection systems were mainly implemented by research or public health groups, such as the Johns Hopkins Applied Physics Laboratory [20,21,35,36], the New York City Department of Health [14], the National Institute of Infectious Diseases in Japan [27], or the Real-Time Outbreak and Disease Surveillance Laboratory at the University of Pittsburg [19]. Data are available the day after the day's sales in those systems. In eight other studies, private marketing companies had automatically aggregated and made available drug sales

data from a sample (2–90%) of pharmacies in a given city or country [18,22,23,31-34,37].

Methods and results by study types

Retrospective studies

Nineteen studies retrospectively compared drug sales data to gold standard reference data of the disease. Details are given in Table 2.

Reference data of the disease included medical case reports [20-22], diagnostic registries of microbiological laboratories [16,24,28,29], hospital admission or discharge data [18-22,27], or clinical emergency department data [14-16]. The selection of indicator drugs in these studies was based on the literature or expert opinion. For example, Edge et al. [17] selected all anti-nauseant and antidiarrheal OTC drugs for gastrointestinal surveillance. In Stirling et al. [26], pharmacists determined which common antidiarrheal drugs they would report.

Two methods were commonly used to compare drug-sale and diagnostic data time series: correlation analysis and signal detection comparison (Table 2). Ten studies used cross-correlation function to measure the similarity of two curves and to determine the time lag at which the correlation between the datasets is maximized. Cross-correlation is a standard method to determine the time delay between two signals. In three studies, only correlation between the time series was examined without analyzing time-lagged relationship. Six studies used aberration detection methods to evaluate whether and by how long the date of signal detection by drug sales precedes the signal based on diagnostic data. The signal definition for aberration detection was based on either a simple threshold to define alerts [15] or more complex algorithms such as the Serfling method [14], ARIMA models [19], the simple moving average method (MA), the cumulative sum method (CUSUM) [11,16], or the exponentially weighted moving average (EWMA) [18]. These studies assessed the performance in terms of sensitivity, specificity and timeliness of disease outbreak detection. Five other studies [13,24-26,29] only evaluated whether drug sales showed a significant increase during a known epidemic period.

Twelve of 14 studies evaluating OTC sales retrospectively found significant correlations or a significant increase in drug sales [14-18,20-22,24-26,29]. Only two studies didn't found any consistent correlation. For example, Das et al. [14] found a poor correlation between OTC antidiarrheal drug sales and emergency department visits for diarrhea in New York City, with an r^2 of 0.24. They found however an increase in sales during a known outbreak of norovirus. OTC drug sales preceded clinical data in three of eight studies that analyzed timeline correlations [18,20,21]. For example, in Hogan et al. [18], the correlation coefficient between electrolyte sales

Table 1 General characteristics of studies included in a literature review of drug sales data analyses for surveillance of infectious diseases

Author	Study period	Location	Syndrome	Drug data sources	Drug status	Time scale
RETROSPECTIVE STUDIES						
Sugawara et al. [27] 2012	2009/2010 2010/2011	Japan	ILI	Prescription drug purchases from 5275 pharmacies	Prescribed	Week
Polgreen et al. [23] 2011	Jan 1999–Sept 2007	Ohio, USA	GI	IMS Health's Xponent database (Data from retail pharmacies, covering 70% of all prescriptions)	Prescribed	Month
Chopra et al. [12] 2011	April–Dec 2009	Southeastern Michigan, USA	ILI	In-site pharmacies of eight hospitals in Southeastern Michigan	Prescribed	Week
Kirian et al. [19] 2010	July 2003–Dec 2007	San Francisco Bay area	GI	Sales available through the National Retail Data Monitor in three San Francisco Bay Area counties (number of stores reporting per week: 592–837)	OTC	Week
Yoshida et al. [30] 2009	Dec 2006–April 2007	Sakai City, Osaka, Japan	ILI	Questionnaires mailed to 273 pharmacies. 56% responding pharmacies. (456,000 people)	Prescribed	Week
Van den Wijngaard et al. [28] 2008	2001–2003	The Netherlands	ILI	Foundation for pharmaceutical statistics (85% of Dutch pharmacies) (13.8 million people)	Prescribed	Week
Edge et al. [17] 2006	Jan 2001–April 2004	Canada (one province)	GI	One major retailer with 19 locations (12% of the pharmacies in the region)	OTC	Week
Das et al. [14] 2005	Aug 2002–March 2005	New York City	ILI and GI	New York City Department of Health (248 NYC pharmacies, 30% of citywide sales)	OTC	Day
Okhusa et al. [22] 2005	Nov 2003–April 2004	Japan	ILI	Private marketing company (1100 pharmacies, 2% of the pharmacies in Japan)	OTC	Day
Chen et al. [11] 2005	March 2003–July 2004	A rural county in New York State	Pertussis	Medicaid database reimbursement	Prescribed	Day
Edge et al. [16] 2004	Jan–May 2001 (Saskatchewan) March–June 2000 (Ontario)	Battleford (Saskatchewan), Walkerton (Ontario)	GI	One pharmacy in Battleford, one pharmacy in Walkerton	OTC	Week
Magruder et al. [21] 2004	2001–2003	Maryland-Washington-Virginia	ILI	Johns Hopkins Applied Physics laboratory (300 drugstores in the Maryland-Washington-Virginia area)	OTC	Day

Table 1 General characteristics of studies included in a literature review of drug sales data analyses for surveillance of infectious diseases (Continued)

Couturier et al. [13] 2004	2001–2003	Paris and five french regions	Syphilis	A centralized wholesaler supplying all French private pharmacies	Prescribed	Month
Hogan et al. [18] 2003	1998–2001	Pennsylvania, Indiana, Utah	ILI and GI	Information Resources, Inc., private company (90% market share in all the region)	OTC	Week
Magruder [20] 2003	2001–2002	Maryland-Washington-Virginia	ILI	Johns Hopkins APL (300 drugstores in the Maryland-Washington-Virginia area)	OTC	Day
Davies et al. [15] 2003	1998/1999, 1999/2000, 2000/2001	Nottingham city	ILI	Retailers (Boots the chemist, 30% of market share) in the hospital area, and a pharmaceutical company (Reckitt Benckiser) at national level	OTC	Week
Stirling et al. [26] 2001	Jan–May 2001	North Battleford, Saskatchewan	GI	Three pharmacies	OTC	Week
Proctor et al. [24] 1998	1993	Milwaukee (Wisconsin, USA)	GI	One pharmacy	OTC	Day
Rodman et al. [25] 1997	1993(Milwaukee)/1994 (Las Vegas)/1996 (Collingwood, Kelowna, Cranbrook)	Milwaukee (Wisconsin); Collingwood (Ontario), Kelowna, Cranbrook (British Columbia), Las Vegas(Nevada)	GI	One pharmacy in Milwaukee, 3 pharmacies in Collingwood, 10 to 12 pharmacies in Kelowna and Cranbrook.	OTC	Month
Welliver et al. [29] 1979	1976/1977	Los Angeles	ILI	One large supermarket chain	OTC	Week
DRUG SELECTION STUDIES						
Pelat et al. [33] 2010	2000–2009	France	GI	IMS Health (7,500 pharmacies in 2000–13,300 in 2008, 59% of the pharmacies in France, covering all continental France)	OTC and prescribed	Week
Cami et al. [31] 2009	Jul 2003- Dec 2006	Region around the city of Houston, TX, USA	ILI	AC Nielsen Corp (pharmacy sales in a region around Houston)	OTC	Week
Wallstrom et al. [34] 2007	2002–2004	Pennsylvania, USA	ILI	AC Nielsen Corp (pharmacy sales in western Pennsylvania)	OTC	Week
Li et al. [32] 2005	Jan 1998- Dec 2001	Pennsylvania, USA	ILI and GI	Information Resources, Inc., private company (90% market share in the region)	OTC	Week
Magruder et al. [21] 2004	2001–2003	Maryland-Washington-Virginia	ILI	Johns Hopkins APL (300 drugstores in the Maryland-Washington-Virginia area)	OTC	Day

Table 1 General characteristics of studies included in a literature review of drug sales data analyses for surveillance of infectious diseases (Continued)

PREDICTION STUDIES							
Vergu et al. [37]	2006	2000-2004	France	ILI	IMS Health (11,000 pharmacies, 50% of all pharmacies in France)	OTC and prescribed	Week
Najmi et al. [36]	2005	2001-2003	Maryland-Washington-Virginia	ILI	Johns Hopkins APL (300 drugstores in the Maryland-Washington-Virginia area)	OTC	Day
Najmi et al. [35]	2004	2001-2002	Maryland-Washington-Virginia	ILI	Johns Hopkins APL (300 drugstores in the Maryland-Washington-Virginia area)	OTC	Day

Abbreviations: *ILI* Influenza-like illness, *GI* Gastrointestinal, *OTC* Over-the-counter drugs.

Table 2 Methodology and results of the retrospective studies included in a literature review of drug sales data analyses for surveillance of infectious diseases

Author	Syndrome	Reference data sources	Drugs selected	Statistical methods	Correlation strength	Correlation timeliness	Detection sensitivity	Detection specificity	Detection timeliness
STUDIES ON PRESCRIBED DRUGS									
Sugawara et al. [27] 2012	ILI	Influenza cases from 5000 hospitals and clinics	Drugs against influenza virus : oseltamivir, zanamivir, laninamivir	Correlation	Pearson corr. coeff. r = 0.992 for 2009/10 and r = 0.972 for 2010/11 (p < 0.001)	-	-	-	-
Polgreen et al. [23] 2011	GI	Hospitalizations with diagnosis of Clostridium Difficile Infections	Oral vancomycin	Cross-correlation, Regression model	Increase in Clostridium Difficile Infections associated with increase in vancomycin use	-	-	-	-
Chopra et al. [12] 2011	ILI	Cases of influenza reported from nine sentinel healthcare providers	Oseltamivir	Correlation	Spearman corr.coeff. r = 0.46 (p < 0.003) Peaks occurred at the same time	-	-	-	-
Yoshida et al. [30] 2009	ILI	28 sentinel surveillance sites of influenza in Sahai City (clinics and hospitals)	Oseltamivir and Zanamivir	Correlation	Pearson corr.coeff. r = 0.954	-	-	-	-
Van den Wijngaard et al. [28] 2008	ILI	Respiratory pathogen diagnosis in laboratory registries (Influenza A, B, RSV, enterovirus, S. pneumoniae..)	Drugs for respiratory infectious diseases (7 ATC classes)	Graphical comparison, Correlation, Linear regression model	Pearson corr. coeff. r = 0.60 for Influenza A, r = 0.58 for RSV, r = 0.60 for S. pneumoniae, r = 0.39 for influenza B (p < 0.05) 80% of variation explained by respiratory pathogens	2 weeks earlier until 1 week later	-	-	-
Author									
Syndrome									
Reference data sources									
Drugs selected									
Statistical methods									
Correlation strength									
Correlation timeliness									
Detection sensitivity									
Detection specificity									
Detection timeliness									
Chen et al. [11] 2005	Pertussis	Reported cases of pertussis from the NYS department of Health	Macrolide antibiotics	CUSUM	-	-	100% The signal was indicator of pertussis outbreak	100%	Not early warning
Couturier et al. [13] 2004	Syphilis	Reported cases of syphilis from hospitals, physicians, sexually transmitted disease clinics.	Benzylpenicillin benzathine 2.4 MU	Descriptive analysis	Similar trend (+22% increase in Paris, +10% in the 5 regions)	Similar trend	-	-	-
STUDIES ON OTC DRUGS									
Kirian et al. [19] 2010	GI	Cases of gastrointestinal diseases from County Health Department and detected GI outbreaks	Diarrheal remedies (based on common use)	Cross-correlation, Regression ARIMA	No significant correlation between sales and GI cases counts; outbreak counts.	-	Not sensitive 4%-14%	Specific 97%-100%	-

Table 2 Methodology and results of the retrospective studies included in a literature review of drug sales data analyses for surveillance of infectious diseases (Continued)

Author	Syndrome	Reference data sources	Drugs selected	Statistical methods	Correlation strength	Correlation timeliness	Detection sensitivity	Detection specificity	Detection timeliness
Edge et al. [17] 2006	GI	Counts of GI cases due to bacteria, parasites, and viruses	Antinauseant and antidiarrheal products	Cross-correlation	Temporal patterns of OTC and Norovirus activity were similar Pearson $r^2 = 0.44$	No delay	-	-	-
Okhusa et al. [22] 2005	ILI	Reporting of patients with ILI (hospitals, clinics, physicians)	Most common treatments ILI	Cross-correlation, Prediction model, Peak comparison	Significant correlation between sales and influenza activity	Sales do not determine influenza in advance	-	-	-
Das et al. [14] 2005	ILI	Emergency department in New York City (ratio of ILI syndrome visits/other syndrome visits)	A cold medication selected statistically from a group of 400 cold medications (ratio ILI/analgesics drugs sales)	Cross-correlation, Serfling method, Graphical comparison	High correlation Pearson $r^2 = 0.60$ ($p < 0.001$)	No lead time	Sensitive (data not reported)	Not specific (not reported)	Not earlier warning than reference data
Das et al. [14] 2005	GI	Emergency department in New York City (ratio of GI syndrome visits/other syndrome visits)	Common antidiarrheal drugs (ratio GI/analgesic drug sales)	Cross-correlation, Graphical comparison	Low correlation Pearson $r^2 = 0.24$ ($p < 0.005$) Similar increases during the fall (norovirus) and influenza peak. Increase in ED GI visits during late winter (rotavirus), but no increase in drug sales.	-	Less sensitive than ED system	-	-
Edge et al. [16] 2004	GI	Emergency room visits for acute GI, number of GI cases from case series investigations (waterborne outbreak)	Saskatchewan: four commonly used antidiarrheals and antinauseants Ontario: 12 products (antidiarrheal, antinauseant, rehydration products)	Graphical comparison (Ontario, Saskatchewan), CUSUM, moving average (Ontario)	Trends of OTC products comparable to the outbreak epidemic curve (Saskatchewan, Ontario)	-	100% exceeded threshold during the outbreak period (Ontario)	100%	Not earlier
Magruder et al. [21] 2004	ILI	Outpatient insurance-claim diagnoses for acute respiratory conditions, from 13,000 clinics and doctors' offices	Remedies for treating influenza (common use)	Cross-correlation	Seasonal trend: Pearson r (between 0.95 and 0.99)	1-3 week lead	-	-	-
					Non-seasonal trend: Pearson r (between 0.25 and 0.75)	No repeatable lead time			

Table 2 Methodology and results of the retrospective studies included in a literature review of drug sales data analyses for surveillance of infectious diseases
 (Continued)

Author	Syndrome	Reference data sources	Drugs selected	Statistical methods	Correlation strength	Correlation timeliness	Detection sensitivity	Detection specificity	Detection timeliness
Hogan et al. [18] 2003	ILI and GI	Hospital-discharge diagnoses of respiratory and diarrheal disease in children (for all hospitals in Pennsylvania, in Utah, and 95% of Indiana).	Electrolyte products	Cross-correlation, EWMA	Pearson $r = 0.90$ (95% CI, 0.87–0.93)	Sales preceded diagnoses by 1.7 weeks (95% CI, 0.50–2.9)	100%	100%	Electrolyte sales preceded detection from diagnoses by an average of 2.4 weeks (95% CI, 0.1–4.8) Detection earlier in 12/18 outbreaks
Magruder [20] 2003	ILI	Outpatient insurance-claim diagnoses for acute respiratory conditions	Cold remedies: 622 products (then grouped in categories by an expert in pharmaco-epidemiology)	Cross-correlation	Pearson $r = 0.9$	Mean lead times of 2.8 days	-	-	-
Davies et al. [15] 2003	ILI	Emergency admission data from Nottingham City Hospital NHS Trust.	Cold and flu remedies (cold, cough, decongestant, throat preparation)	Correlation, Peak comparison, Threshold detection method	National and local sales positively correlated with admissions in 98/99 and 99/00, not 00/01	-	100%(for local sales)	100% (for local sales)	Rate of local sales exceed threshold of 1000 units per week 2 weeks prior to peak in emergency admissions
Stirling et al. [26] 2001	GI	Telephone survey from a sample of households: number of persons with diarrheal symptoms and/or with stool specimen positive to <i>C. parvum</i> oocysts. (waterborne outbreak)	Common antidiarrheal (determined by each pharmacist)	Descriptive analysis	A fivefold increase in sales during the epidemic period	-	-	-	-
Proctor et al. [24] 1998	GI	Comparison with eight sources (laboratory confirmed cases of <i>Cryptosporidium</i> , clinically defined cases) (waterborne outbreak)	Antidiarrheal: Imodium, Pepto Bismol, Kaopectate	Descriptive analysis	Significant increase in drug sales during epidemic period	-	-	-	-

Table 2 Methodology and results of the retrospective studies included in a literature review of drug sales data analyses for surveillance of infectious diseases (Continued)

Rodman et al. [25] 1997	GI	Cases of cryptosporidiosis (5 waterborne outbreaks)	Antidiarrheal drugs	Descriptive analysis	Milwaukee: increased 20 fold; Las Vegas: no data; Collingwood: increased in 2 of 3 stores; Kelowna: increased 3 fold; Cranbrook: increased	-	-
Welliver et al. [29] 1979	ILI	Laboratory count of influenza B	Children's aspirin, adult antipyretics, cold remedies	Determination of the % of sales increase, peak comparison	Sales of cold remedies averaged 185% above the baseline value during the peak influenza activity	-	-

Abbreviations: *ILI* Influenza-like illness, *GI* Gastrointestinal, *RSV* Respiratory syncytial virus, *ATC* Anatomical Therapeutic Chemical classification system, *CUSUM* Cumulative sum control chart, *OTC* Over-the-counter drugs, *EWMA* Exponentially Weighted Moving Average.

and hospital diagnoses of respiratory and diarrheal illness was 0.90 (95% CI, 0.87–0.93) when drug sales were assumed to precede clinical diagnosis data by 1.7 weeks. Outbreaks were detected with 100% sensitivity and specificity in 3 of 5 studies that analyzed signal detection [15,16,18]. Drug sales data provided an earlier outbreak signal in two of them [15,18]. In Davies et al. [15], the rate of cough/cold sales exceeded a threshold of 1000 units per week two weeks before the peak in emergency department admissions during three consecutive winters. In Hogan et al. [18], detection from electrolytes sales occurred an average 2,4 weeks earlier than detection from hospital diagnoses of respiratory and diarrheal diseases.

Six of the seven studies that focused on prescribed drugs found strong correlations ($r = 0.46–0.99$) with clinical reference data or a significant increase in drug sales, without lead time however. The other study [11] showed that the CUSUM signal generated for prescriptions for macrolide antibiotics was linked to a pertussis outbreak in a county of New York State.

No association was observed between the type of reference data and the time lags observed.

Drug selection studies

An important challenge for drug-sales-based surveillance is identifying relevant indicator drug groups to monitor diseases. Five retrieved articles addressed this question. Characteristics of the studies are described in Table 3.

Two studies [21,34] developed methods to find homogeneous groups of OTC products. The authors used unsupervised clustering algorithms for aggregating OTC products in groups sharing similar sales histories. For example, Magruder et al. [21] first assigned OTC products for respiratory diseases to subgroups qualitatively based on indication, dose form, and age group. A stepwise hierarchical clustering algorithm was then used to form

categories sharing a similar sales history, leading to a set of 16 product categories.

In two studies [31,33], the authors developed procedures to identify the drugs correlating with disease incidence. Clusters were formed specifically for a particular disease. In Pelat et al. [33], a hierarchical clustering procedure was applied to the time series of all therapeutic classes and the acute diarrhea incidence rate reported by a network of general practitioners. Four therapeutic classes were found to cluster with diarrhea incidence and an algorithm based on the selected drugs allowed the detection of epidemics with a sensibility of 100%, a specificity of 95% and a timeliness of 1.7 weeks before official alerts.

Prediction studies

In three studies [35-37], the authors developed models to predict clinical data based on drug sales data.

Vergu et al. [37] used a Poisson regression model on selected OTC sales to forecast influenza-like illness (ILI) incidence as recorded by a sentinel network of general practitioners. The forecast at the national level 1–3 weeks ahead showed a strong correlation with observed ILI incidence ($r = 0.85–0.96$).

Najmi et al. [35] used least mean square filtering methods to estimate the incidence of emergency room consultations for respiratory conditions from past and present sales of groups of cold-remedy sales. In a later article [36], they succeeded in extending the estimation algorithm for predicting increases in clinical data several days in advance.

Discussion

The evidence gathered in this systematic literature review suggests that drug sales data analysis can be a useful tool

Table 3 Methodology and results of drug selection studies included in a literature review of drug sales data analysis for surveillance of infectious diseases

Author	Disease	Method	Results of the algorithm evaluation
Pelat et al. [33] 2010	GI	Hierarchical clustering procedure ,CUSUM	Identification of 4 therapeutic classes relevant to gastroenteritis outbreak detection. Detection performance of a multiple voter algorithm: sensibility 100%, specificity 95%, timeliness 1.7 weeks.
Cami et al. [31] 2009	ILI	Aggregate mining algorithm	Identification of product categories with outbreak detection performance superior to predefined categories and more strongly correlated with the disease data.
Wallstrom et al. [34] 2007	ILI	Unsupervised time-series clustering algorithm	Distinction between OTC products for allergy and OTC products for influenza symptoms
Li et al. [32] 2005	ILI/GI	Canonical correlation analysis	Identification of eight diagnoses that have strong association with electrolyte sales ($r = 0.96$)
Magruder et al. [21] 2004	ILI	Unsupervised stepwise clustering algorithm	Identification of 16 OTC product groups with similar historical trends

Abbreviations: GI Gastrointestinal, CUSUM Cumulative sum control chart, ILI Influenza-like illness, OTC Over-the-counter drugs.

for surveillance of acute respiratory and gastrointestinal infections.

As could be expected, prescribed drug sales data were strongly correlated with clinical case reporting. No lead time was observed, which is consistent with the fact that patients purchase drugs after seeing a healthcare professional. Analysis of prescribed drug sales data may nevertheless have an additional utility for epidemic detection, as these data might be available with a shorter delay than clinical surveillance data [27].

A high correlation between OTC drug sales data and reference surveillance data were found in almost all the retrospective studies. Several studies also showed that OTC drug sales can serve as an early indicator of disease epidemics. Patients may buy nonprescription drugs during the early phase of illness when they become symptomatic, before consulting a health practitioner [38]. A surveillance system based on drug data should ideally detect all the outbreaks, rapidly, with a low false alert rate. However, few studies in the review determined the sensitivity and specificity of the outbreak detection and those aspects should be analyzed in more details in future studies.

Surveillance based on OTC drug sales could be particularly relevant for diseases whose prodromal phase persists for several days before the onset of more severe symptoms. For example, the early stages of dengue fever symptoms are nonspecific (fever, headache, myalgia, arthralgia, etc.) [39]. The occurrence of grouped cases could trigger an excess of nonspecific drug sales over baseline levels, which in turn could provide an early warning of outbreak in an endemic area.

Results from drug selection studies showed that it is possible to identify groups of products strongly associated with incidence data, which can then be used to predict future trends in clinical data and help public health authorities to prepare response planning. Such product selection procedures, however, depend on the existence of large clinical surveillance databases of the diseases concerned.

Similarly, the validity of drug sales data analysis has been evaluated mainly for two disease groups, respiratory and gastrointestinal illness, for which clinical reference data, used as the gold standard, are readily available. Pertussis and syphilis have been evaluated in only one study each, and still require further confirmation. The concept of drug-based surveillance therefore needs to be validated for other infectious diseases.

All the studies were conducted in developed countries or area. Surveillance based on drug sales data requires electronic information systems for routine data analysis. Besides, its implementation requires that the population has access to the health care system and mainly buy drugs in pharmacies. This could limit the use of drug based surveillance systems in developing countries.

By improving the timeliness of epidemic detection compared to clinical data and giving information from a larger part of the population, drug sales data can be an additional source of information for already monitored diseases. Besides, drug sales data analysis could have its greatest value in the surveillance of diseases for which clinical surveillance is cumbersome and costly, or where substantial under-reporting is suspected. To confirm the selected drug group as a valid proxy of disease, clinical surveillance may be conducted for a defined period in a representative population. Examples of diseases for which this would be useful are typically varicella, urinary infections, allergies/asthma, and parasitic diseases.

Ideally, the drugs to be monitored should be specific to the disease and widely used to treat it in order to maximize the sensitivity of the signal. For example, benzylpenicillin benzathine 2.4 MUI is the quasi exclusive treatment for syphilis infection [13] and is a good candidate. In contrast, the treatment of measles is mostly symptomatic without a specific drug, which makes this disease unattractive for this approach. Another limitation applies to diseases that are usually treated in hospitals or specialized centers, such as tuberculosis.

Surveillance based on drug sales, may not be appropriate to accurately estimate incidence of diseases, as the source population size is not precisely known. Moreover, it may be difficult to link the number of drug packages sold to the number of patients with disease. However, the method is very efficient to determine temporal dynamics of a situation and to detect abnormal phenomena. Surveillance based on drug sales is therefore well adapted to diseases with seasonal variations such as norovirus gastroenteritis, influenza and other infectious respiratory agents, or community outbreaks (foodborne illnesses, waterborne illnesses, hepatitis A, etc.).

Drug sales can be influenced by store promotions, sales period (holidays, weekends), and the media. Also, we do not know whether people buy medications to treat a disease they currently have or a disease they fear they may have in the near future. For example, during the media coverage of avian influenza A (H5N1) in the US, an increase in antiviral medications sales was observed [40], which corresponded to stockpiling behavior of the population.

Health-seeking behaviors also vary by demographic, social, cultural, and economic factors. A survey [41] in Canada analyzed the healthcare-seeking behaviors of 351 patients with acute gastroenteritis. They found significant differences (patient age and sex) between the patients who used OTC drugs and those who did not. Consequently, factors that prompt self-medication should also be taken into account. The usefulness of drug sales based surveillance is also dependent on the available resources and the organization of the health care system. OTC drug sales

surveillance is for example less relevant in countries where reimbursement rate are high and patients mainly get prescribed drugs.

Population mobility, particularly in tourist areas, may lead to an increase in remedy sales, which could wrongly be interpreted as a disease outbreak. Inversely, patients with high geographical mobility may not be included in the region of study and lead to an underestimation of the magnitude of an epidemic.

Despite some limitations, routine collection and analysis of drug sales data are likely to be developed in the coming years. Many automated surveillance systems that collect drug data the day after the sales have been implemented in the last decade [14,19,20,27]. They allow a rapid assessment of the public health situation. Early detection of outbreaks allows public health authorities to set up epidemic investigations and control measures sooner. Most studies included in this review were published after the year 2000, with their number increasing recently. They illustrate the need for improved surveillance systems, evidenced by recent public health crises (e.g., anthrax in 2001, the SARS outbreak in 2003, the A/H1N1 influenza pandemic in 2009, etc.). Drug sales data present indeed many advantages in terms of public health surveillance. Data can be obtained in a real-time manner and usually cover a large portion of the population. Data collection may be exhaustive, without selection of specific sales, and allows the simultaneous monitoring of a large number of diseases, especially new or emerging diseases.

Although non-specific, drug sales data are directly linked to patients' health conditions. Drug sales data are therefore more specific than other syndromic surveillance data, such as tracking search patterns on the web and are likely to reflect more accurately disease activity in the population. Moreover, it should be noted that alternative sources of data for disease surveillance are currently under development. Healthcare management databases that can provide exhaustive information on drug consumption and diagnosis, as the Dossier Médical Personnel [42] in France, are promising tools for disease surveillance.

Our review may be affected by a publication bias since studies unable to show correlations between drug sales and reference data may have been less published. In addition, selection bias may have occurred in the studies. Indeed, some studies in the review were based on a limited number of pharmacies and/or a limited study period (e.g. less than one year). Language bias may exist as we were not able to identify studies published in languages other than English and French. The review focused on the temporal dynamics of infectious disease; consequently, further analyses are required to determine the capacity of these systems to efficiently monitor other aspects of infectious diseases such as spatial spreading.

Conclusion

This review suggests that the analysis of drug sales data is a promising method for surveillance and outbreak detection of infectious diseases. It has the potential to trigger an outbreak alert earlier than most surveillance systems. However, the main challenges consist in the appropriate selection of indicator drug groups and the validation of this approach for diseases for which no or poor quality clinical surveillance data exists. The usefulness of the approach also depends on the available resources and the organization of the health care system. Drug sales databases with real-time or near real-time data transmission are available in several countries; future studies should be encouraged to expand their use on other infectious diseases.

Competing interests

Mathilde Pivette prepares a doctoral thesis under the French framework "CIFRE" (Industrial contract for training through research; www.anrt.asso.fr), in partnership with the company Celtipharm (www.celtipharm.org). The other authors declare they have no competing interests.

Authors' contributions

All authors contributed to the study's design. MP and JM carried out the literature search and reviewed articles. MP drafted the manuscript. All authors interpreted the results, revised and approved the final manuscript.

Authors' information

Mathilde Pivette has a Doctor of Pharmacy degree and a Master of Public Health. She is a PhD candidate in epidemiology at the French School of Public Health. Her research interest lies in the analysis of drug sales data for disease surveillance.

Acknowledgements

This research was funded by Celtipharm (Vannes, France) a company specialized in the real time collection and statistical processing of healthcare data (www.celtipharm.org – www.openhealth.fr), through a doctoral thesis contract for Mathilde Pivette.

Author details

¹EHESP French School of Public Health, Sorbonne Paris Cité, Rennes, France. ²Université Paris Descartes, MAP5, Paris, France. ³Celtipharm, Vannes, France. ⁴Institut Pasteur, Emerging Diseases Epidemiology Unit, Paris, France. ⁵Aix Marseille Université, IRD French Institute of Research for Development, EHESP French School of Public Health, UMR_D 190 "Emergence des Pathologies Virales", Marseille, France.

Received: 24 May 2013 Accepted: 31 October 2014

Published online: 18 November 2014

References

1. Henning KJ: What is syndromic surveillance? *MMWR Morb Mortal Wkly Rep* 2004, **53**(Suppl):5–11.
2. CDC: Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CEC working group. *MMWR Morb Mortal Wkly Rep* 2004, **53**(RR-5):1–11.
3. Berger M, Shiau R, Weintraub JM: Review of syndromic surveillance: implications for waterborne disease detection. *J Epidemiol Community Health* 2006, **60**:543–550.
4. Kara EO, Elliot AJ, Bagnall H, Foord DGF, Praisner R, Osman H, Smith GE, Olowokure B: Absenteeism in schools during the 2009 influenza A(H1N1) pandemic: a useful tool for early detection of influenza activity in the community? *Epidemiol Infect* 2011, **140**:1–9.
5. Rolland E, Moore KM, Robinson VA, McGuinness D: Using Ontario's "Telehealth" health telephone helpline as an early-warning system: a study protocol. *BMC Health Serv Res* 2006, **6**:10.

6. Rodman JS, Frost F, Jakubowski W: **Using nurse hot line calls for disease surveillance.** *Emerg Infect Dis* 1998, **4**:329–332.
7. Lewis MD, Pavlin JA, Mansfield JL, O'Brien S, Boomsma LG, Elbert Y, Kelley PW: **Disease outbreak detection system using syndromic data in the greater Washington DC area.** *Am J Prev Med* 2002, **23**:180–186.
8. Josseran L, Fouillet A, Caillère N, Brun-Ney D, Illeff D, Brucker G, Medeiros H, Astagneau P: **Assessment of a syndromic surveillance system based on morbidity data: results from the Oscour network during a heat wave.** *PLoS One* 2010, **5**:e11984.
9. Mostashari F, Fine A, Das D, Adams J, Layton M: **Use of ambulance dispatch data as an early warning system for communitywide influenzalike illness, New York City.** *J Urban Heal* 2003, **80**(2 Suppl 1):i43–i49.
10. Moher D, Liberati A, Tetzlaff J, Altman DG: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** *BMJ* 2009, **339**:b2535.
11. Chen J-H, Schmit K, Chang H, Herlihy E, Miller J, Smith P: **Use of Medicaid prescription data for syndromic surveillance—New York.** *MMWR Morb Mortal Wkly Rep* 2005, **54**(Suppl):31–34.
12. Chopra T, Binienda J, Mohammed M, Shyamraj R, Long P, Bach D, Carlton C, Peters S, Lephart P, Alangaden G, Dhar S, Marchaim D, Schreiber M, Kaye KS: **A practical method for surveillance of novel H1N1 influenza using automated hospital data.** *Infect Control Hosp Epidemiol* 2011, **32**:700–702.
13. Couturier E, Michel A, Janier M, Dupin N, Semaille C: **Syphilis surveillance in France, 2000–2003.** *Euro Surveill* 2004, **9**:8–10.
14. Das D, Metzger K, Heffernan R, Balter S, Weiss D, Mostashari F: **Monitoring over-the-counter medication sales for early detection of disease outbreaks—New York City.** *MMWR Morb Mortal Wkly Rep* 2005, **54**(Suppl):41–46.
15. Davies GR, Finch RG: **Sales of over-the-counter remedies as an early warning system for winter bed crises.** *Clin Microbiol Infect* 2003, **9**:858–863.
16. Edge VL, Pollari F, Lim G, Aramini J, Sockett P, Martin SW, Wilson J, Ellis A: **Syndromic surveillance of gastrointestinal illness using pharmacy over-the-counter sales. a retrospective study of waterborne outbreaks in Saskatchewan and Ontario.** *Can J Public Heal* 2004, **95**:446–450.
17. Edge VL, Pollari F, Ng LK, Michel P, McEwen S a, Wilson JB, Jerrett M, Sockett PN, Martin SW: **Syndromic surveillance of norovirus using over-the-counter sales of medications related to gastrointestinal illness.** *Can J Infect Dis Med Microbiol* 2006, **17**:235–241.
18. Hogan WR, Tsui F-C, Ivanov O, Gesteland P, Grannis S, Overhage JM: **Detection of pediatric respiratory and diarrheal outbreaks from sales of over-the-counter electrolyte products.** *J Am Med Informatics Assoc* 2003, **10**:555–562.
19. Kirian ML, Weintraub JM: **Prediction of gastrointestinal disease with over-the-counter diarrheal remedy sales records in the San Francisco Bay Area.** *BMC Med Inform Decis Mak* 2010, **10**:39.
20. Magruder SF: **Evaluation of over-the-counter pharmaceutical sales as a possible early warning indicator of human disease.** *John Hopkins Univ Appl Phys Lab Tech Dig* 2003, **24**:349–353.
21. Magruder SF, Lewis SH, Najmi A, Florio E: **Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance.** *MMWR Morb Mortal Wkly Rep* 2004, **53**(Suppl):117–122.
22. Ohkusa Y, Shigematsu M, Taniguchi K, Okabe N: **Experimental surveillance using data on sales of over-the-counter medications—Japan, November 2003–April 2004.** *MMWR Morb Mortal Wkly Rep* 2005, **54** Suppl:47–52.
23. Polgreen PM, Yang M, Kuntz JL, Laxminarayan R, Cavanaugh JE: **Using oral vancomycin prescriptions as a proxy measure for Clostridium difficile infections: a spatial and time series analysis.** *Infect Control Hosp Epidemiol* 2011, **32**:723–726.
24. Proctor ME, Blair KA, Davis JP: **Surveillance data for waterborne illness detection: an assessment following a massive waterborne outbreak of Cryptosporidium infection.** *Epidemiol Infect* 1998, **120**:43–54.
25. Rodman JS, Frost F, Davis-Burchat L: **Pharmaceutical sales; a method for disease surveillance?** *J Environ Health* 1997, **60**:8–14.
26. Stirling R, Aramini J, Ellis A, Lim G, Meyers R, Fleury M, Werker D: **Waterborne cryptosporidiosis outbreak.** *Canada Commun Dis Rep* 2001, **27**:185–192.
27. Sugawara T, Ohkusa Y, Ibuka Y, Kawahara H, Taniguchi K, Okabe N: **Real-time prescription surveillance and its application to monitoring seasonal influenza activity in Japan.** *J Med Internet Res* 2012, **14**:e14.
28. Van den Wijngaard C, van Asten L, van Pelt W, Nagelkerke NJD, Verheij R, de Neeling AJ, Dekkers A, van der Sande MAB, van Vliet H, Koopmans MPG: **Validation of syndromic surveillance for respiratory pathogen activity.** *Emerg Infect Dis* 2008, **14**:917–925.
29. Welliver RC, Cherry JD, Boyer KM, Deseda-Tous JE, Krause PJ, Dudley JP, Murray RA, Wingert W, Champion JG, Freeman G: **Sales of nonprescription cold remedies: a unique method of influenza surveillance.** *Pediatr Res* 1979, **13**:1015–1017.
30. Yoshida M, Matsui T, Ohkusa Y, Kobayashi J, Ohyama T, Sugawara T, Yasui Y, Tachibana T, Okabe N: **Seasonal influenza surveillance using prescription data for anti-influenza medications.** *Jpn J Infect Dis* 2009, **62**:233–235.
31. Cami A, Wallstrom GL, Fowlkes AL, Panozzo CA, Hogan WR: **Mining aggregates of over-the-counter products for syndromic surveillance.** *Pattern Recognit Lett* 2009, **30**:255–266.
32. Li R, Wallstrom GL, Hogan WR: **A multivariate procedure for identifying correlations between diagnoses and over-the-counter products from historical datasets.** In *AMIA Annu Symp Proc.* 2005:450–454.
33. Pelat C, Boëlle P-Y, Turbelin C, Lambert B, Valleron A-J: **A method for selecting and monitoring medication sales for surveillance of gastroenteritis.** *Pharmacoepidemiol Drug Saf* 2010, **19**:1009–1018.
34. Wallstrom GL, Hogan WR: **Unsupervised clustering of over-the-counter healthcare products into product categories.** *J Biomed Inform* 2007, **40**:642–648.
35. Najmi A-H, Magruder SF: **Estimation of hospital emergency room data using otc pharmaceutical sales and least mean square filters.** *BMC Med Inform Decis Mak* 2004, **5**:1–5.
36. Najmi A-H, Magruder SF: **An adaptive prediction and detection algorithm for multistream syndromic surveillance.** *BMC Med Inform Decis Mak* 2005, **5**:33.
37. Vergu E, Grais RF, Sarter H, Fagot J-P, Lambert B, Valleron A-J, Flahault A: **Medication sales and syndromic surveillance, France.** *Emerg Infect Dis* 2006, **12**:416–421.
38. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, Pavlin J, Gesteland P, Treadwell T, Koski E, Hutwagner L, Buckeridge DL, Aller RD, Grannis S: **Implementing syndromic surveillance: a practical guide informed by the early experience.** *J Am Med Informatics Assoc* 2004, **11**:141–150.
39. Meynard J-B, Chaudet H, Texier G, Ardillon V, Ravachol F, Deparis X, Jefferson H, Dussart P, Morvan J, Boutin J-P: **Value of syndromic surveillance within the Armed Forces for early warning during a dengue fever outbreak in French Guiana in 2006.** *BMC Med Inform Decis Mak* 2008, **8**:29.
40. Miller J, Schmit K, Duncan P, Waters C: **Increased antiviral medication sales before the 2005–06 influenza season—New York City.** *MMWR Morb Mortal Wkly Rep* 2006, **55**:277–279.
41. Frosst GO, Majowicz SE, Edge VL: **Factors associated with the use of over-the-counter medications in cases of acute gastroenteritis in Hamilton, Ontario.** *Can J public Heal* 2006, **97**:489–493.
42. DMP- Dossier Médical Personnel. www.dmp.gouv.fr.

doi:10.1186/s12879-014-0604-2

Cite this article as: Pivette et al.: Drug sales data analysis for outbreak detection of infectious diseases: a systematic literature review. *BMC Infectious Diseases* 2014 **14**:604.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



IV. Surveillance de la gastro-entérite en France à partir des ventes de médicaments

Cette étude sur la détection des épidémies de gastro-entérites en France à partir des ventes de médicaments en pharmacies a fait l'objet d'une publication scientifique en septembre 2014 dans la revue *Epidemics* :

M. Pivette, J.E. Mueller, P. Crépey, and A. Bar-Hen, "Surveillance of gastrointestinal disease in France using drug sales data," *Epidemics*, vol. 8, pp. 1–8, Sep. 2014.

1) Résumé

Le suivi des ventes de médicaments en pharmacies est un outil de surveillance qui s'est développé au cours des dix dernières années pour la surveillance des maladies infectieuses et la détection d'épidémies. Notre revue de la littérature a montré que les ventes de médicaments étaient de bons indicateurs des maladies grippales et gastro-intestinales et que les médicaments sans ordonnance avaient le potentiel de détecter précocement les épidémies (Pivette et al. 2014a). La gastro-entérite est une maladie infectieuse causée par des virus ou bactéries et qui présente une épidémie saisonnière (1 700 000 cas estimés en moyenne par saison hivernale pour la période 2008 à 2013 en France). Une détection précoce des épidémies est ainsi nécessaire pour pouvoir mettre en place des mesures de prévention et ajuster les messages de communication. L'objectif de notre étude était d'évaluer la validité du suivi des ventes de médicaments pour la détection des épidémies saisonnières de gastro-entérites en France et de déterminer leur intérêt pour la détection précoce des épidémies à un niveau national et régional. Pour la période 2008-2013, nous avons comparé les séries temporelles des ventes de médicaments prescrits et non prescrits indiqués pour la gastro-entérite et du nombre de cas de diarrhées aiguës vus en consultation.

Les ventes de médicaments sont issues de l'échantillon stratifié de 3004 pharmacies en France Métropolitaine, développé par l'entreprise *Celtipharm*. Nous avons sélectionné des médicaments susceptibles d'être prescrits ou achetés en cas gastro-entérites en nous basant sur la classification EphMRA des médicaments. La sélection finale contenait 8 classes de médicaments, ce qui correspond à 256 produits. Nous avons différencié les médicaments prescrits et ceux achetés sans ordonnance. Les données de référence proviennent d'un réseau

de 1300 médecins qui rapportent chaque semaine le nombre de cas de diarrhées aiguës vus en consultation (French GPs Sentinelles Network). Les données sont extrapolées au niveau national. Nous avons analysé la corrélation entre le nombre de boîtes de médicaments vendues et les cas de gastro-entérites par semaine. Le modèle de régression périodique de Serfling a été utilisé pour déterminer les périodes épidémiques dans les séries temporelles de ventes de médicaments. Nous avons testé différents modèles de détection afin de sélectionner celui qui permettait d'optimiser la sensibilité, le taux de fausses alertes et le délai de détection à partir des médicaments, en prenant comme référence le début des épidémies déterminé à partir des données cliniques (+/- 4 semaines). Les modèles de détection ont été testés avec des données hebdomadaires et journalières. Les modèles de détection ont également été appliqués aux données de ventes régionales afin d'analyser la propagation géographique lors de chaque épidémie. Pour évaluer la capacité de détection en temps réel, nous avons développé une méthode de détection prospective et l'avons évalué pour la détection de l'épidémie pendant la saison hivernale 2012/2013.

De fortes corrélations étaient observées entre le nombre de boîtes vendues et le nombre de cas. Cette corrélation était maximale avec un délai d'une semaine pour les médicaments non-prescrits (coeff corr 0,77). Pour les médicaments prescrits, la corrélation était forte (coeff corr 0,89) et sans délai. Le pic hivernal avec les médicaments non-prescrits était en moyenne 1,2 semaine plus tôt que celui observé avec les données cliniques. Les modèles de détection basés sur les ventes de médicaments ont permis de détecter les cinq épidémies saisonnières, sans fausse alerte. La date de détection avec les médicaments prescrits était similaire à celle des données de référence. En revanche, la détection à partir des médicaments non-prescrits était en moyenne 2,25 semaines plus tôt. L'analyse prospective à partir des médicaments non-prescrits a montré que l'épidémie de 2012/2013 aurait pu être détectée trois semaines en avance en calibrant le modèle sur les quatre années précédentes. L'analyse régionale a montré que certaines régions étaient en période épidémique avant même la détection nationale et qu'il s'agissait d'un outil intéressant pour analyser la diffusion géographique de la maladie.

Les résultats confirment l'intérêt du suivi des ventes de médicaments pour la détection d'épidémies saisonnières de gastro-entérites en France. Les ventes de médicament non prescrits, achetés en pharmacie lors des premiers symptômes, permettent de détecter de façon précoce le début des périodes épidémiques. Les données sont obtenues automatiquement à J+1, permettant la détection en temps réel des épidémies et l'évaluation rapide de la situation. Ce type de surveillance épidémiologique ne vise qu'un seul des objectifs de la surveillance

épidémiologique qui est de pouvoir alerter en cas de suspicion d'épidémie. Le résultat recherché est de pouvoir proposer le plus rapidement possible des mesures sanitaires pour bloquer la propagation de la maladie. Ce modèle a montré son utilité pour détecter le début des épidémies mais des études complémentaires seraient nécessaires pour voir son intérêt pour détecter la fin d'une épidémie ou l'intensité de son pic. L'analyse a porté sur les épidémies saisonnières de fortes intensités pour lesquelles des données de référence existent ; pour valider cette approche pour la détection d'épidémies localisées il serait nécessaire de mettre en place des investigations sur le terrain.

Notre étude montre que les ventes de médicaments sont des outils performants pour la détection des épidémies hivernales de gastro-entérites en France et une distinction doit être faite entre les médicaments prescrits et les médicaments non-prescrits qui eux permettent une détection précoce. Ces méthodes de détection à partir des médicaments pourraient être développées pour d'autres maladies infectieuses en France.

2) Article



Surveillance of gastrointestinal disease in France using drug sales data



Mathilde Pivette^{a,b,c,*}, Judith E. Mueller^{a,d}, Pascal Crépey^{a,e}, Avner Bar-Hen^b

^a EHESP French School of Public Health, Sorbonne Paris Cité, Rennes, France

^b Université Paris Descartes, MAP5, Paris, France

^c Celtipharm, Vannes, France

^d Institut Pasteur, Emerging Diseases Epidemiology Unit, Paris, France

^e Aix Marseille Université, IRD French Institute of Research for Development, EHESP French School of Public Health, UMR.D 190 "Emergence des Pathologies Virales", Marseille, France

ARTICLE INFO

Article history:

Received 18 December 2013

Received in revised form 2 May 2014

Accepted 6 May 2014

Available online 20 May 2014

Keywords:

Population surveillance

Disease outbreaks

Drug utilization

Epidemiology

Gastroenteritis

ABSTRACT

Drug sales data have increasingly been used for disease surveillance during recent years. Our objective was to assess the value of drug sales data as an operational early detection tool for gastroenteritis epidemics at national and regional level in France. For the period 2008–2013, we compared temporal trends of drug sales for the treatment of gastroenteritis with trends of cases reported by a Sentinel Network of general practitioners. We benchmarked detection models to select the one with the best sensitivity, false alert proportion and timeliness, and developed a prospective framework to assess the operational performance of the system. Drug sales data allowed the detection of seasonal gastrointestinal epidemics occurring in winter with a distinction between prescribed and non-prescribed drugs. Sales of non-prescribed drugs allowed epidemic detection on average 2.25 weeks earlier than Sentinel data. These results confirm the value of drug sales data for real-time monitoring of gastroenteritis epidemic activity.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-SA license (<http://creativecommons.org/licenses/by-nc-sa/3.0/>).

Introduction

Drug sales data have increasingly been used in recent years for outbreak detection and surveillance of acute diseases, like respiratory (Das et al., 2005; Davies and Finch, 2003; Magruder et al., 2004; Ohkusa et al., 2005; Vergu et al., 2006; Magruder, 2003) or diarrheal diseases (Das et al., 2005; Edge et al., 2004, 2006; Stirling et al., 2001; Proctor et al., 1998; Hogan et al., 2003; Pelat et al., 2010; Kirian and Weintraub, 2010; Rodman et al., 1997). Although generally unspecific, drug sales data can provide information within short delays, capture a large part of the population and may detect changes in the population health status (Henning, 2004).

Gastroenteritis is a highly infectious disease caused by viruses, bacteria and parasites. Symptoms are commonly vomiting and diarrhea. Early detection of the start of gastroenteritis epidemics could limit their spreading, since preventative measures could be implemented and public health messages delivered in a timely manner. Previous studies have shown that medications sales

for gastroenteritis are a good proxy of gastroenteritis incidence (Edge et al., 2004, 2006; Stirling et al., 2001; Proctor et al., 1998) and that outbreak onsets could be detected retrospectively 1.75 weeks–2.4 weeks before reference data (Hogan et al., 2003; Pelat et al., 2010). The objective of the present study was to assess the value of drug sales data analysis as an operational epidemic detection tool for gastroenteritis seasonal epidemics in France. We focused on gastroenteritis as it is a public health issue for which early and valid detection is of particular interest and because surveillance data of the disease, reported by a Sentinel Network of physicians, are available in France (French GPs Sentinelles Network). We refined analyses by using non-prescribed and prescribed drug sales separately, as over-the-counter (OTC) drugs have shown to be early indicators of diseases in previous studies (Davies and Finch, 2003; Magruder et al., 2004; Vergu et al., 2006; Magruder, 2003; Hogan et al., 2003; Najmi and Magruder, 2005). We first retrospectively assessed the correlation between medication sales and GP-reported cases of gastroenteritis in France. We performed a benchmark of detection models to select the one with the best sensitivity, false alert proportion and early detection, and we present a method which can be used prospectively for detecting seasonal epidemics based on weekly or daily drug sales data at both national and regional scale.

* Corresponding author at: EHESP, Avenue du Professeur Léon Bernard, 35043 Rennes, France. Tel.: +33 6 48412957.

E-mail address: mathilde.pivette@ehesp.fr (M. Pivette).

Methods

Data sources

Daily drug sales data were extracted from a stratified sample of community pharmacies on the French metropolitan territory set up by the company Celtipharm, Vannes, France (Celtipharm). The sample size increased from 1086 pharmacies in 2008 to 3004 (13% of the total pharmacies) in 2013. Real time sales data of these pharmacies are automatically and continuously collected, every day, since 2007. Thanks to a constant up-to-date database of all the 22,458 French active community pharmacies, Celtipharm built a stratification on sale revenue (6 levels for global revenue and 4 levels per type of sales: prescribed drugs, OTC, and other type of sales), localisation (5 geographic areas) and sales area (5 types, from rural to densely urban). Each stratum has a minimum of 30 pharmacies or is merged with neighbouring strata. Sampling rates per strata are computed with the Neyman optimal allocation algorithm (Kish, 1965). Extrapolations from this sample have been validated with data from drug manufacturers who distribute their products to pharmacies. This database has already been used in previous studies (Crépey et al., 2013a,b).

Reference surveillance data were obtained from the French Sentinel Network of physicians (French GPs Sentinelles Network; Valleron et al., 1986; Flahault et al., 2006). The network is constituted of 1300 volunteer general practitioners working in metropolitan France (2% of the total GPs). They transmit each week data from their patient consultations on 8 health indicators via Internet connections. Cases of acute diarrhea are reported since 1990. A case of acute diarrhea is defined as a patient having at least three daily, soft or watery stools in the past 14 days. Weekly data are extrapolated to the national level and are accessible online (French GPs Sentinelles Network). In France, surveillance of gastroenteritis is also based on hospital emergency data and laboratory data in case of investigation of localized outbreaks (school, nursing home. . .) (Invs French Institute for Public Health Surveillance). However, we used Sentinel Network data as reference because they are more representative of the general population.

Indicator drugs selection

We selected groups of products likely to be prescribed or bought for gastroenteritis. Pharmaceutical products are aggregated into classes according to the Anatomical Classification of the European Pharmaceutical Marketing Research Association (EPHMR) based on their anatomical site of action, main indication, therapeutic use and composition. We selected the following medication classes (class code): intestinal anti-infectives antidiarrheals (A07A), intestinal adsorbents antidiarrheals (A07B), antidiarrheal microorganisms (A07F), motility inhibitors (A07H), other antidiarrheals (A07X), other antiemetics and antinauseants (A04A9). Drugs for motion sickness were excluded from the antiemetic group (A04A9). We also selected two classes of parapharmaceutical products: oral rehydration solutions and alimentary products for vomiting and diarrhea commonly used for children. The total selection contains 8 classes, corresponding to 256 products (Fig. S1).

Data

The study period lasted from January 7, 2008 (week 2 2008) to June 30, 2013 (week 26 2013). We used the weekly number of cases per 100,000 inhabitants. For drugs, we used the weekly and daily number of boxes sold per 100,000 inhabitants at the national and regional level. We differentiated products prescribed by a health practitioner and products purchased without prescription.

Correlation analysis

Cross-correlation analyses were performed to measure the similarity of the time series at different time lags and the dates of peak in each time series were compared. Since reference data were obtained on a weekly basis, correlation analysis and peak comparisons were performed on weekly data.

Epidemics detection method

To identify epidemic periods in drug sales data series, we applied a Serfling-type periodic regression model, a widely-used method for detecting outbreaks for diseases with a seasonal background pattern (Costagliola et al., 1991). The method was first proposed by Serfling (1963) and more recently implemented by Pelat et al. (2007) in a software application. The drug data series are assumed to be composed of a periodic baseline level and epidemic periods. We first removed a highest N percentile of the observations (N is the pruning value) in the training period (e.g. the highest 30%). A model is then fitted on the remaining data to estimate a non-epidemic baseline level. We modeled the baseline level with a linear trend and commonly used periodicities of 12, 6 and 3 months. We defined the threshold as the upper bound of the prediction interval (PI) of the non-epidemic level (Pelat et al., 2007). An epidemic is defined when a defined number of observations are above the threshold, depending on the detection rule. The onset of an epidemic corresponds to the first observation above the threshold.

Primary and secondary analyses were performed using weeks and days, respectively, as time unit. For daily analyses, we removed Sundays and public holidays, as most pharmacies in France are closed those days. We also removed one day before and after public holidays, as sales are higher than expected. We added a weekly periodicity in the periodic regression model based on the sales of each weekday to take into account the day-of-the-week effect (Table S1) described in other studies (Ohkusa et al., 2005; Magruder, 2003). The model equations are given in appendix (Text S1).

To identify the best detection model, we tested different key parameters. We excluded between 15% and 40% of the highest observations in the training period. Thresholds were defined by taking the upper limit of the 90%, 95% and 99% prediction interval. We tested detection rules of 1 or 2 consecutive weeks above the threshold to define an epidemic for weekly analyses, and detection rules of 2–8 consecutive epidemic days for daily analyses.

Epidemics detection evaluation

The reference epidemic periods were the epidemic periods at the national level as they had been defined and published by the Sentinel Network, using periodic regression models on acute diarrhea incidence rates (French GPs Sentinelles Network). We defined a “detection window” of 4 weeks before and 4 weeks after the onset of epidemics defined by the Sentinel Network. The choice of a wide detection window will enable the detection of very early signals of epidemics, until 4 weeks before reference data.

A true positive alert was defined as an epidemic onset computed from drug data occurring in the detection window. A false positive alert was defined as an epidemic onset computed from drug data occurring outside the detection window. The sensitivity represents the proportion of seasonal epidemics that are detected by the drug-based detection system, i.e. if at least one positive alert occurred in the detection window. False alert proportion was defined as the ratio of false positive alerts to the total number of alerts. We focused only on the start of the epidemics, regardless of their duration. Timeliness is defined as the time difference between the first true positive alert and the start of epidemic period from Sentinel Network. Timeliness is undefined when no alert is emitted in the

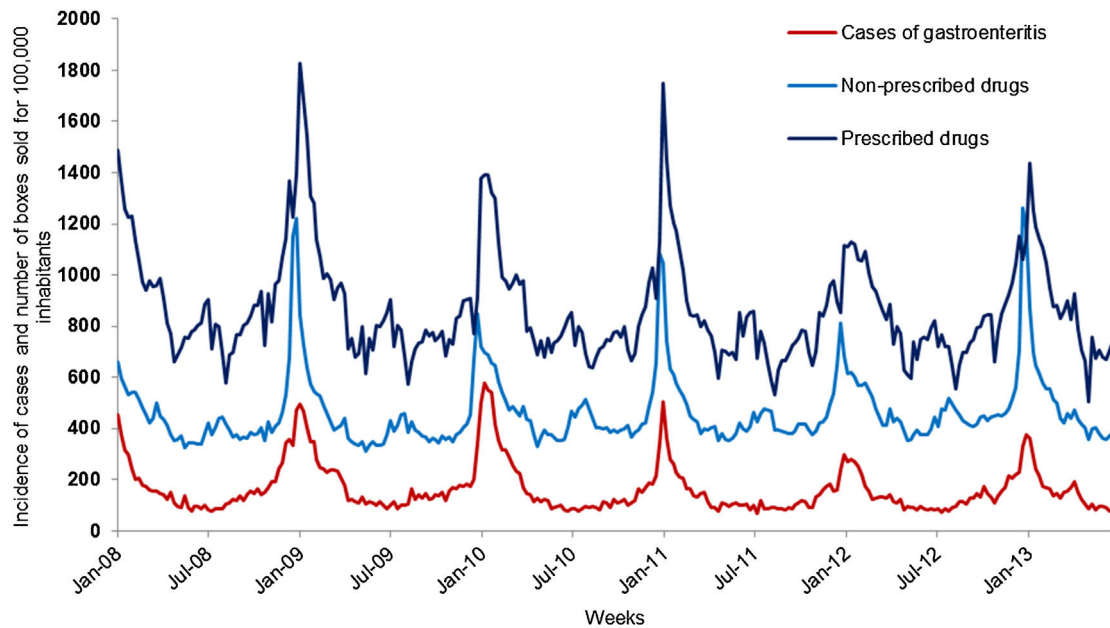


Fig. 1. Weekly sales of prescribed and non-prescribed drugs for gastroenteritis (in number of boxes sold for 100,000 inhabitants), and incidence of gastroenteritis obtained from a Sentinel Network of general practitioners (in number of cases for 100,000 inhabitants). Metropolitan France, calendar week 2 of 2008 to week 26 in 2013.

detection window. Note that a true positive alert can occur at maximum 4 weeks before the detection from Sentinel Network, given our detection window. For daily data, timeliness was calculated using as reference the first day (Monday) of the first epidemic week from Sentinel Network, as epidemic detection from Sentinel is only based on weekly data.

We calculated sensitivity, false alert proportion and timeliness for the different models by varying the key parameters (pruning value, PI and detection rule). The objective was to find the set of parameters allowing first the highest sensitivity, then the lowest false alert proportion, and we finally kept, among that selection, the parameters allowing the earliest detection.

Geographical analyses

To evaluate the surveillance tool at regional level and to describe geographical dynamics of gastroenteritis, we applied the selected periodic regression models on regional time series of drug sales data. In each region, every week was defined as epidemic or non-epidemic. We constructed maps representing the weekly epidemic status of the 21 administrative regions of metropolitan France (Corsica excluded).

Prospective detection during 2012/2013

To assess the real time performance of the detection model, we evaluated the epidemic detection performance on the last year (2012–2013) based on weekly and daily non-prescribed drug sales, after fitting a periodic regression model on the four preceding years (2008–2012). We first selected the best detection model for the period 2008–2012, using the methodology previously described. The selected model is then extrapolated to forecast an epidemic threshold for the following year (2012–2013). We evaluated sensitivity, false alert proportion and timeliness of the detection in 2012–2013.

All analyses were performed on R Statistical software (www.r-project.org) and Microsoft Office Excel 2010. Maps were prepared with the R statistical software.

Results

Descriptive analysis

The five main peaks in the data series of diarrhea incidence rates were also observed in the sale rates of prescribed and non-prescribed drugs (Fig. 1), and in each of the eight individual drug classes (Fig. S2). All main peaks occurred during winter months. Milder increases during summer months were also observed in drug sales series, but not with diarrheal cases. This summer increase appeared among all classes, being most pronounced for sales of oral rehydration solutions and motility inhibitors. Weekly incidences ranged from 66 to 577 cases per 100,000 inhabitants. Weekly drug boxes sold for 100,000 inhabitants ranged from 310 to 1260 for non-prescribed drugs and from 503 to 1829 for prescribed drugs. Overall, the number of prescribed drug sales was 1.9 times higher than the number of non-prescribed drug sales (in France, drugs need to be prescribed to be reimbursed by the French Health insurance).

The prescribed drug sales and the number of reported cases were highly correlated, with a maximum correlation coefficient of 0.89 at no time lag. The maximum correlation between non-prescribed drug sales and cases was observed at a time lag of -1 week with a coefficient of 0.77.

During all five epidemic periods, the peak occurred first in non-prescription drug sales and on average 1.2 weeks before the peak in Sentinel cases. Peaks in prescribed drug sales lagged behind Sentinel data peaks by on average 0.7 weeks (Table 1).

Evaluation of epidemic detection performance

Non prescribed drug sales

Weekly data analyses. Analyzing weekly sales of non-prescription drugs, the best detection performance was obtained in a model with a pruning value of 30%, a PI of 99% and with an epidemic period being defined when at least two weeks exceed the threshold (Table 2). In this model, non-prescription drugs allowed detecting all five seasonal epidemics, with no false alert and detection occurred on average 1.6 weeks earlier (median: -2 ; min: -3 ; max:

Table 1
Timing of winter peak for each time series and peak comparison between drug sales data and consultation-based sentinel data. Metropolitan France, week 2 of 2008 to week 26 of 2013.

Winter period	Cases (calendar week of peak)	Non prescribed drugs (calendar week of peak)	Time difference between the peaks in non-prescribed drugs and in Sentinel data (weeks)	Prescribed drugs (calendar week of peak)	Time difference between the peaks in prescribed drugs and in Sentinel data (weeks)
2008/2009	2009 02	2009 01	−1	2009 02	0
2009/2010	2010 02	2009 53	−2	2010 03	+1
2010/2011	2011 01	2010 52	−1	2011 01	0
2011/2012	2012 01	2011 52	−1	2012 03	+2
2012/2013	2013 01	2012 52	−1	2013 02	+1
Mean timeliness (weeks)			−1.2		0.7

Table 2
Sensitivity, false alert proportion and timeliness in detection of gastroenteritis from weekly non-prescribed drug sales. Metropolitan France, week 2 of 2008 to week 26 of 2013. The best detection model was obtained with a pruning value of 30%, and a prediction interval of 99% (gray shade).

PI (%)	Pruning value (%)	Sensitivity	False alert proportion	Mean timeliness (weeks)	Median timeliness (weeks)
90	15	1	0.17	−1.6	−2
90	20	1	0.17	−2.0	−2
90	25	1	0.29	−2.2	−3
90	30	1	0.50	−2.4	−3
90	35	1	0.55	−2.4	−3
90	40	0.8	0.67	−2.0	−2
95	15	1	0.17	−1.2	−2
95	20	1	0.17	−1.6	−2
95	25	1	0.17	−2.0	−2
95	30	1	0.38	−2.2	−3
95	35	1	0.44	−2.4	−3
95	40	0.8	0.69	−2.0	−2
99	15	1	0	−1.0	−1
99	20	1	0	−1.2	−2
99	25	1	0	−1.4	−2
99	30	1	0	−1.6	−2
99	35	1	0.38	−2.2	−3
99	40	1	0.44	−2.4	−3

+1). Detection was earlier in the four last seasonal epidemics, with a mean timeliness of 2.25 weeks (Table 3, Fig. 2A).

Daily data analyses. Analyzing daily data, the best detection performance was obtained in a model with a pruning value of 25%, a PI of 90% and with an epidemic period being defined when at least seven consecutive observations were higher than the threshold. The resulting sensitivity was of 100% with no false alert, and the detection occurred on average 12.4 days earlier (median: −16; min: −25; max: +5;). Detection occurred first from non-prescribed drugs in the four last epidemics with a mean timeliness of 16.75 days (Table S2, Fig. S3).

Prescribed drug sales

Weekly data analyses. Analyzing weekly sales of prescribed drugs, the best performing model used a pruning value of 30%, a PI of 95% a detection rule of two consecutive epidemic weeks (Table S3). Sensitivity was 100%, with no false alert, and a detection occurring

0.6 weeks earlier than reference data (median: 0; min: −2; max: +1) (Fig. 2B, Table 3).

Daily data analyses. The best detection performance with daily data was obtained for a pruning value of 40%, a PI of 95% and a detection rule of six days. We obtained a sensitivity of 80% with 20% of false alert, and an average timeliness of +3.6 days (median: 1; min: 0; max: 14). No parameters allowed 100% sensitivity without false alert rate. The lower detection performance can be explained by high inter-week variations in daily prescribed drug sales.

Prospective detection in 2012–2013

Using the best performing models fitted only on the four preceding years 2008–2012 (pruning value of 20% and PI 95% for weekly data; pruning value of 20% and PI 90% for daily data), the epidemic period in 2012–2013 was detected without false alert. Based on weekly and daily data, the epidemic period started, respectively, 3

Table 3
Start of epidemic periods obtained each winter from non-prescribed drugs (based on the selected model with a pruning value of 30% and a PI of 99%), from prescribed drugs (based on the selected model with a pruning value of 30% and a PI of 95%) and timeliness of detection compared to Sentinel data. Metropolitan France, week 2 of 2008 to week 26 of 2013.

Year	Sentinel Network epidemic period (calendar week of start)	Non prescribed drugs epidemic period (calendar week of start)	Timeliness of detection from non prescribed drugs compared to Sentinel data (weeks)	Prescribed drugs epidemic period (calendar week of start)	Timeliness of detection from prescribed drugs compared to Sentinel data (weeks)
2008/2009	2008 49	2008 50	+1	2008 47	−2
2009/2010	2009 53	2009 52	−1	2010 01	+1
2010/2011	2010 52	2010 50	−2	2010 52	0
2011/2012	2012 01	2011 50	−3	2012 01	0
2012/2013	2012 52	2012 49	−3	2012 50	−2
Mean timeliness			−1.6		−0.6
Median timeliness			−2		0

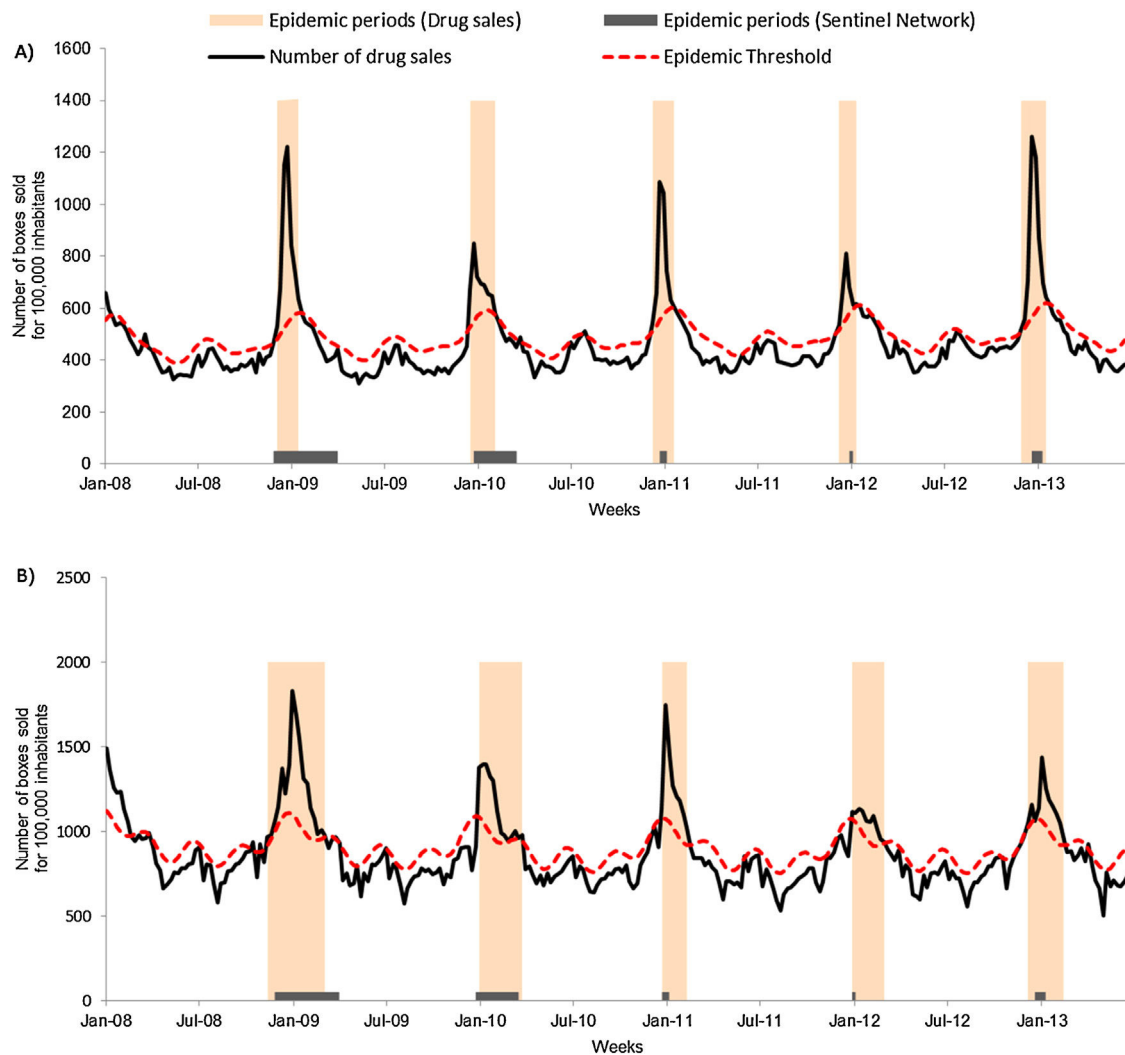


Fig. 2. Weekly sales of non-prescribed drugs (A) and prescribed drugs (B) with the corresponding epidemic periods defined by the selected detection models. Epidemic periods defined by the Sentinel Network are represented with gridded surfaces.

weeks and 19 days earlier than the period defined in the Sentinel Network (Fig. 3). As 2 consecutive epidemic weeks and 7 consecutive epidemic days were necessary to confirm an epidemic, the start of the seasonal epidemic could have been declared on December 13 using daily data and on December 16 using weekly data, whereas declaration based on the Sentinel Network (French GPs Sentinelles Network) was done on January 7.

Geographical analyses

In Fig. 4, we represented the regions affected by gastroenteritis epidemic every week for each winter period. During the four winters 2009–2013, we observed that some regions were detected as epidemic even before the national detection from drug sales. For example, based on non-prescribed drug sales analysis in 2011–2012 and 2012–2013, epidemic was first detected in the northeast regions of France and one week before the national detection. We did not use reference data at regional level to evaluate the detection performance. Indeed, the sample of sentinel physicians is relatively low and heterogeneous, which implies too much variability at a regional level to use them as reference.

Discussion

Our study presents a method for real-time detection of gastroenteritis epidemics based on drug sales data. Our results suggest that non-prescribed drug sales allow epidemic detection substantially earlier than consultation-based sentinel data.

Over-the-counter (OTC) sales of anti-diarrheals, anti-nauseants and rehydration products have proven to be good indicators of gastroenteritis illness in several previous studies (Edge et al., 2004, 2006; Stirling et al., 2001; Proctor et al., 1998; Hogan et al., 2003; Rodman et al., 1997) but only two studies have shown the potential of drug sales for early detection of gastroenteritis compared to traditional surveillance data. In Hogan et al. (2003), detection from OTC electrolytes sales occurred an average 2.4 weeks earlier than detection from hospital diagnoses of respiratory and diarrheal pediatric diseases in the eighteen outbreaks studied. In Pelat et al. (2010), four therapeutic classes correlating with diarrhea incidence were selected and an algorithm based on the selected classes allowed the detection of seasonal gastrointestinal epidemics in France with a sensitivity of 100%, a specificity of 95% and a timeliness of 1.7 weeks before official alerts. However, this study did not make any distinction between OTC and prescription

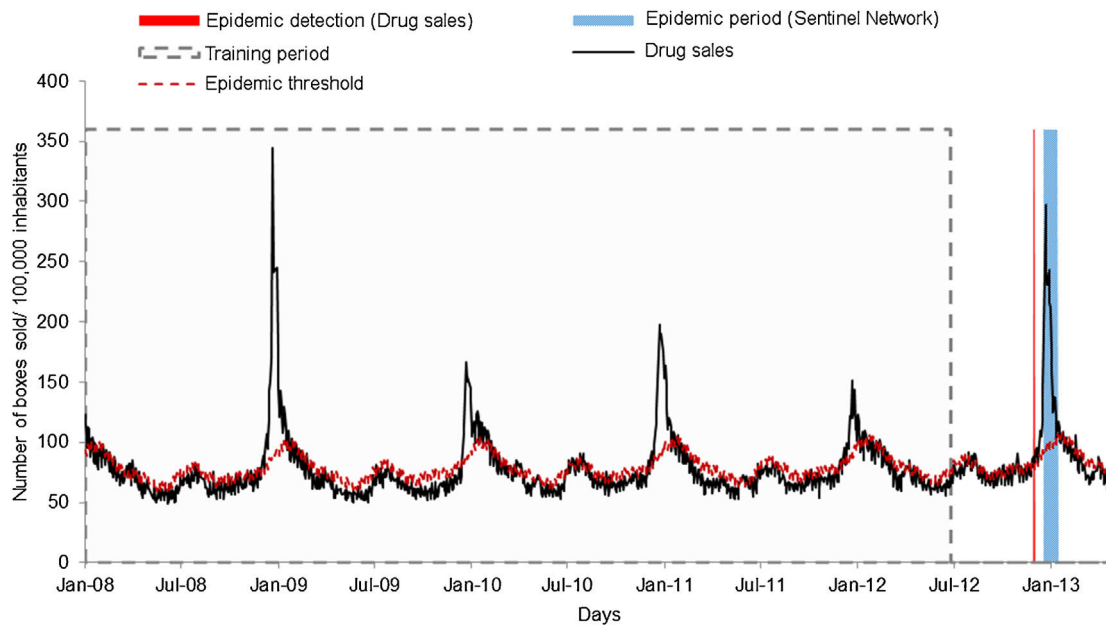


Fig. 3. Prospective epidemic detection in 2012–2013 using daily non-prescribed drug sales data. Metropolitan France, week 2 of 2008 to week 26 of 2013. The vertical red line indicates the epidemic detection from drug sales data. The 2012–2013 epidemic was detected nineteen days earlier than Sentinel Network detection. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

drugs and used only a retrospective analysis framework that do not allow conclusion on usefulness in real-time surveillance.

Contrary to published studies that focused on OTC drug sales (Das et al., 2005; Edge et al., 2004, 2006; Stirling et al., 2001; Proctor et al., 1998; Hogan et al., 2003; Kirian and Weintraub, 2010; Rodman et al., 1997), which can be purchased with or without a prescription, we analyzed products which were actually purchased without a prescription. Our analysis therefore reflects more precisely self-medication behavior of the population, which appears to allow more timely detection of epidemics.

We found that the correlation between non-prescribed drugs and cases was highest with a one week lag-time, indicating a different timing in patient behavior. Non-prescription drug sales peaked earlier than the number of cases every year. Non-prescribed drug sales analysis allowed detecting all seasonal epidemics without false alert and the detection was earlier than reference data in four out of five epidemics with a mean timeliness of 2.25 weeks. Using day as time unit, the detection was on average 16.75 days earlier. This result suggests that non-prescribed drug sales

have the potential to be a sensitive, specific and timely indicator of seasonal gastroenteritis epidemics. It seems that patients buy non-prescribed drugs during the early phase of illness when they become symptomatic, before or instead consulting a physician. The detection occurred later with non-prescribed drugs only during the first year (2008–2009) of the study period; however, the sample of pharmacies was smaller at the beginning of the study period and may partly explain the differences observed with the following years.

By contrast, we found that prescribed drug sales provided relatively little benefits in terms of early detection compared to Sentinel reference surveillance data (on average 0.6 weeks). We found high correlations between prescribed drug sales and the number of reported cases, indicating that both time-series peak approximately at the same time. This result confirmed the adequacy between the two data sources and is explained by the fact that patients buy drugs just after seeing a physician. Prescribed drug sales can therefore be used as a source of information to monitor gastroenteritis disease and is a good indicator of medical consultations.

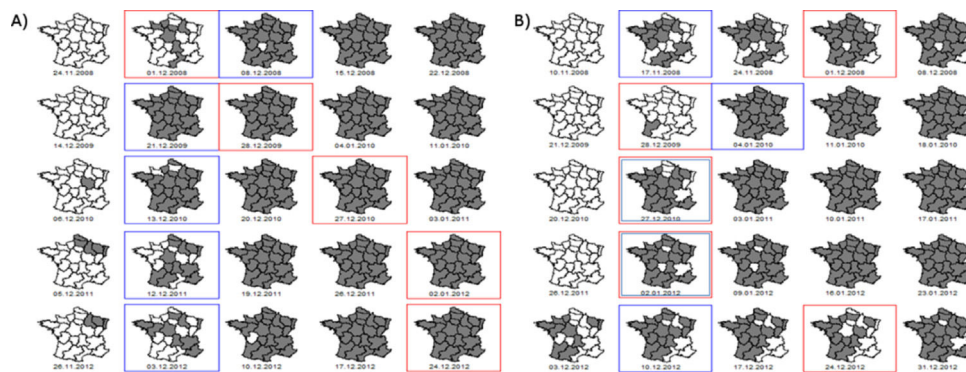


Fig. 4. Weekly epidemic regions defined using non-prescribed (A) and prescribed drugs (B) are represented in black for each winter period. The detection weeks obtained from drug sales analysis at the national level are framed by a blue line, the detection weeks from Sentinel Network data at the national level are framed by a red line. Metropolitan France (Corsica excluded), 2008–2013 winter seasons. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Data sales in our database were obtained on a daily basis, whereas a majority of surveillance studies were based on weekly data. For non-prescribed drugs, the dates of epidemic detection did not differ between analysis of daily and weekly data. The main advantage of daily sales data analysis is the number of observations above the threshold necessary to declare an epidemic period. Contrary to weekly data for which two consecutive epidemic weeks are needed, only seven consecutive epidemic working days are necessary for epidemic detection based on daily data. Epidemic periods can therefore be declared earlier using daily data and public health response can be delivered more rapidly.

In our data collection system, data were available for analysis one day after the day of sale. This near-real-time collection allowed the use of prospective analyses to detect epidemics. Our prospective analysis of the 2012–2013 season showed that non-prescribed drug sales data would have allowed the detection of the epidemic start three weeks or 19 days earlier than the GP based sentinel system, and the warning would have been sent 25 days earlier than official alerts. This confirms the potential of non-prescribed drugs for timely real-time detection and suggests that the method may be used to detect seasonal epidemics in advance. This gain in timeliness may be of interest for the short term planning of healthcare resources (hospital wards and staff management) and for timing more effectively public health campaigns aimed at limiting the spread of the disease, but it may be of an even greater interest in case of emergence or re-emergence of specific pathogens (e.g. *E. coli*).

One limitation of our study is the selection of category of products, which was based on their indications. These products do not all have the same specificity for gastroenteritis and may have others indications. However, analysis of products categories is a simple method that limits the impact of individual drugs changes due to marketing-induced behavior or others causes of variability (e.g. sales promotion, advertising, out-of-stock issues, Hogan and Wagner, 2006). Besides, some drugs that may be used in case of gastroenteritis, such as analgesics or antipyretics, were not included as there were not specific enough of the disease. We could also not differentiate drugs bought for stockpiling in prevention of the seasonal gastroenteritis from the one immediately used. However, sudden and massive stockpiling by the population is unlikely to occur without a triggering public event, which could then be taken into account in the analysis. It is also likely that a medication purchase is more difficult to trigger than a web search query, which makes our analysis more robust than web queries based systems (Olson et al., 2013; Pelat et al., 2009). The observed seasonality cannot be due to bias on data collection. Indeed, control series, such as hormonal contraceptives, presented a stable number of sales during the study period (data not shown). The use of medications may vary by demographic factors, including age and gender (Frosst et al., 2006), and only a part of the population with acute gastroenteritis consults or buys medications. A telephone-based survey (Van Cauteren et al., 2012), conducted in France on 260 patients reporting acute gastroenteritis, found that only 30.9% of the participants purchased prescription-based medications and that 8.9% purchased OTC drugs. However, underestimation is also an issue in other types of surveillance system, as the same study showed that only 33.4% of the patients consulted a physician. The size of the source population is not precisely known. Therefore, analysis of drug sales may not be appropriate to accurately estimate the incidence of gastro-enteritis in a population but is still a relevant tool to determine dynamics of the disease.

We focused on the high-intensity national seasonal gastroenteritis epidemics for which reference clinical data were available and reliable. The detection windows of 4 weeks before and after the epidemic onset from reference data did not allow the detection of localized outbreaks occurring during other periods of the year.

For example, we observed increases in drug data sales during summer months; one hypothesis is that the increase reflects localized gastroenteritis outbreaks in summer. They are of low-intensity, localized in smaller geographical region, and not currently captured by the national network of general practitioners. Considering the large number of pharmacies in the sample (3004) even at a regional level (around 200 on average), we can assume that analyzing drug sales can be a useful alternative way to monitor epidemic spread at a local level. Data sales are available at the pharmacy level, and sales increase in a localized geographical area could therefore be detected rapidly. However to validate this approach, field investigations have to be conducted and field data are needed (laboratory data, GP consultations data, telephone-based data in the study area) to compare drug sales data to the number of cases.

We could not currently validate our regional analysis with reference data since the French GP-based sentinel network is too heterogeneous to provide reliable local estimates on a weekly basis. However, our regional analysis suggests that local epidemics could be detected even earlier than the national one, which could be another way to further improve public health alerts.

Drug sales data present many advantages in term of public health surveillance. Data can be obtained in almost real-time, which allows rapid assessment of a situation. Efforts and cost to automatically collect sales data are lower compared to other types of surveillance system. Finally, data collection is exhaustive for each pharmacy, enabling the monitoring of a large number of diseases at the same time.

Conclusion

Drug sales data contains information necessary to detect gastrointestinal epidemics occurring in winter but a distinction between prescribed and non-prescribed drugs should be made in order to maximize the timeliness of the detection. Sales of non-prescribed drugs allow epidemic detection on average 2.25 weeks earlier than GP reported clinical data. Our analyses confirmed that drug sales data analysis is a valid and useful tool for gastroenteritis real-time and spatial surveillance. Analysis of drug sales could also be interesting for the surveillance of diseases for which clinical surveillance does not exist, is costly, or poorly reported. Therefore, future drug sales studies should be developed for other infectious diseases.

Acknowledgments

This research was funded by Celtipharm (Vannes, France) a company specialized in the real time collection and statistical processing of healthcare data (www.celtipharm.org – www.openhealth.fr), through a doctoral thesis contract for Mathilde Pivette.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at [doi:10.1016/j.epidem.2014.05.001](https://doi.org/10.1016/j.epidem.2014.05.001).

References

- Celtipharm (Internet). Available from: www.celtipharm.org
- Costagliola, D., Flahault, A., Galinec, D., Gamenn, P., Menares, J., 1991. A routine tool for detection and assessment of epidemics of influenza-like syndromes in France. *Am. J. Public Health* 81 (11), 97–99.
- Crépey, P., Pivette, M., Bar-hen, A., 2013a. Quantitative assessment of preventive behaviors in France during the Fukushima nuclear crisis. *PLoS ONE* 8 (3), 1–5.
- Crépey, P., Pivette, M., Desvarieux, M., 2013b. Potential impact of influenza A/H1N1 pandemic and hand-gels on acute diarrhea epidemic in France. *PLoS ONE* 8 (10), 1–6.

- Das, D., Metzger, K., Heffernan, R., Balter, S., Weiss, D., Mostashari, F., 2005. Monitoring over-the-counter medication sales for early detection of disease outbreaks – New York city. *MMWR – Morb. Mortal. Wkly. Rep.* 54 (August (Suppl.)), 41–46.
- Davies, G.R., Finch, R.G., 2003. Sales of over-the-counter remedies as an early warning system for winter bed crises. *Clin. Microbiol. Infect.* 9 (August (8)), 858–863.
- Edge, V.L., Pollari, F., Lim, G., Aramini, J., Sockett, P., Martin, S.W., et al., 2004. Syndromic surveillance of gastrointestinal illness using pharmacy over-the-counter sales. A retrospective study of waterborne outbreaks in Saskatchewan and Ontario. *Can. J. Public Health* 95 (6), 446–450.
- Edge, V.L., Pollari, F., Ng, L.K., Michel, P., McEwen, S.A., Wilson, J.B., et al., 2006. Syndromic surveillance of norovirus using over-the-counter sales of medications related to gastrointestinal illness. *Can. J. Infect. Dis. Med. Microbiol.* 17 (July (4)), 235–241.
- EPHMRA (Internet). Available from: www.ephmra.org
- Flahault, A., Blanchon, T., Dorléans, Y., Toubiana, L., Vibert, J.F., Valleron, A.J., 2006. Virtual surveillance of communicable diseases: a 20-year experience in France. *Stat. Methods Med. Res.* 15 (October (5)), 413–421.
- French GPs Sentinelles Network (Internet). Available from: <http://websenti.u707.jussieu.fr/sentiweb>
- Frosst, G.O., Majowicz, S.E., Edge, V.L., 2006. Factors associated with the use of over-the-counter medications in cases of acute gastroenteritis in Hamilton, Ontario. *Can. J. Public Health* 97 (6), 489–493.
- Henning, K.J., 2004. What is syndromic surveillance? *MMWR – Morb. Mortal. Wkly. Rep.* 53 (September (Suppl.)), 5–11.
- Hogan, W.R., Wagner, M.M., 2006. Sales of over-the-counter healthcare products. In: Wagner, M.M., Moore, A.W., Aryel, R.M. (Eds.), *Handbook of biosurveillance*. Elsevier Inc., Burlington, MA, pp. 321–331.
- Hogan, W.R., Tsui, F.-C., Ivanov, O., Gesteland, P., Grannis, S., Overhage, J.M., 2003. Detection of pediatric respiratory and diarrheal outbreaks from sales of over-the-counter electrolyte products. *J. Am. Med. Inf. Assoc.* 10 (6), 555–562.
- Invs French Institute for Public Health Surveillance. Surveillance of acute gastroenteritis in France (in French) (Internet). Available from: <http://www.invs.sante.fr/fr/Dossiers-thematiques/Maladies-infectieuses/Risques-infectieux-d-origine-alimentaire/Gastro-enterites-aigues-virales/Contexte-et-dispositif-de-surveillance>
- Kirian, M.L., Weintraub, J.M., 2010. Prediction of gastrointestinal disease with over-the-counter diarrheal remedy sales records in the San Francisco Bay area. *BMC Med. Inform. Decis. Mak.* January (10), 39.
- Kish, J.L., 1965. *Survey Sampling*. Wiley & Sons, New York, NY, pp. 643.
- Magruder, S.F., 2003. Evaluation of over-the-counter pharmaceutical sales as a possible early warning indicator of human disease. *John Hopkins Univ. Appl. Phys. Lab. Tech. Dig.* 24, 349–353.
- Magruder, S.F., Lewis, S.H., Najmi, A., Florio, E., 2004. Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance. *MMWR – Morb. Mortal. Wkly. Rep.* 53 (September (Suppl.)), 117–122.
- Najmi, A.-H., Magruder, S.F., 2005. An adaptive prediction and detection algorithm for multistream syndromic surveillance. *BMC Med. Inform. Decis. Mak.* 5 (January), 33.
- Ohkusa, Y., Shigematsu, M., Taniguchi, K., Okabe, N., 2005. Experimental surveillance using data on sales of over-the-counter medications – Japan, November 2003–April 2004. *MMWR – Morb. Mortal. Wkly. Rep.* 54 (August (Suppl.)), 47–52.
- Olson, D.R., Konty, K.J., Paladini, M., Viboud, C., Simonsen, L., 2013. Reassessing Google flu trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput. Biol.* 9 (10).
- Pelat, C., Boëlle, P.-Y., Cowling, B.J., Carrat, F., Flahault, A., Ansart, S., et al., 2007. Online detection and quantification of epidemics. *BMC Med. Inform. Decis. Mak.* 7 (January), 29.
- Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A., Valleron, A.-J., 2009. More diseases tracked by using Google trends. *Emerg. Infect. Dis.* 15 (8), 1327–1328.
- Pelat, C., Boëlle, P.-Y., Turbelin, C., Lambert, B., Valleron, A.-J., 2010. A method for selecting and monitoring medication sales for surveillance of gastroenteritis. *Pharmacoepidemiol. Drug Saf.* 19 (October (10)), 1009–1018.
- Proctor, M.E., Blair, K.A., Davis, J.P., 1998. Surveillance data for waterborne illness detection: an assessment following a massive waterborne outbreak of *Cryptosporidium* infection. *Epidemiol. Infect.* 120 (February (1)), 43–54.
- Rodman, J.S., Frost, F., Davis-Burchat, L., 1997. Pharmaceutical sales; a method for disease surveillance? *J. Environ. Health*, 8–14.
- Serfling, R.E., 1963. Methods for current statistical analysis of excess pneumonia–influenza deaths. *Public Health Rep.* 78 (6), 494–506.
- Stirling, R., Aramini, J., Ellis, A., Lim, G., Meyers, R., Fleury, M., et al., 2001. Waterborne cryptosporidiosis outbreak. *Can. Commun. Dis. Rep.* 27 (22), 185–192.
- Valleron, A.J., Bouvet, E., Garnerin, P., Ménarès, J., Heard, I., Letrait, S., et al., 1986. A computer network for the surveillance of communicable diseases: the French experiment. *Am. J. Public Health* 76 (November (11)), 1289–1292.
- Van Cauteren, D., De Valk, H., Vaux, S., Le Strat, Y., Vaillant, V., 2012. Burden of acute gastroenteritis and healthcare-seeking behaviour in France: a population-based study. *Epidemiol. Infect.* 140 (April (4)), 697–705.
- Vergu, E., Grais, R.F., Sarter, H., Fagot, J.-P., Lambert, B., Valleron, A.-J., et al., 2006. Medication sales and syndromic surveillance, France. *Emerg. Infect. Dis.* 12 (March (3)), 416–421.

3) Autres méthodes de détection

Dans notre article, nous avons fait le choix d'utiliser la méthode de régression périodique de *Serfling* pour détecter les épidémies dans les séries temporelles des ventes de médicaments. D'autres méthodes de détection existent dans la littérature pour détecter des épidémies à partir de données de surveillance syndromique (Burkom 2003; Jackson et al. 2007). Nous présenterons dans cette partie, de façon non exhaustive, d'autres méthodes de détection utilisées dans la littérature et discuterons leurs limites et avantages pour la détection des épidémies.

Dans notre revue de la littérature, nous avons vu que des méthodes de détection appelées cartes de contrôle (*control charts*) ont été utilisées pour détecter les épidémies à partir des ventes de médicaments (Pivette et al. 2014a). On peut notamment citer les méthodes CUSUM (Edge et al. 2004; Chen et al. 2005) et Exponentially Weighted Moving Average control chart (Hogan et al. 2003). Nous présenterons ces méthodes de cartes de contrôle et les appliquerons à nos séries temporelles de ventes de médicaments pour la gastro-entérite.

Les maladies infectieuses présentent souvent une augmentation saisonnière qu'il est nécessaire de prendre en compte pour estimer la valeur attendue. Il faut alors considérer que les données suivent un niveau de base périodique auquel s'ajoutent parfois des épidémies. Ainsi, la méthode de régression de Serfling (Serfling 1963) utilisée dans notre article a permis de modéliser un niveau de base (sans épidémie) avec la prise en compte d'une tendance, des saisonnalités et une fluctuation aléatoire. Nous présenterons ici deux autres méthodes basées sur les seuils historiques, la méthode de Stroup et la méthode de Farrington, qui permettent également de prendre en compte les saisonnalités.

a) Cartes de contrôle

(i) *La méthode CUSUM*

Principe de la méthode

La méthode CUSUM (Cumulative Sum Control Chart) est une méthode séquentielle permettant de détecter des changements dans une série d'observations. Elle a été initialement proposée pour le contrôle de qualité dans l'industrie, et a ensuite été utilisée pour la surveillance des évènements de santé (Wohl 1977).

A chaque temps, la différence entre la valeur observée et la valeur attendue est calculée. Le CUSUM correspond à la somme cumulée de ces différences sur les temps précédents. Le CUSUM au temps t correspond donc à la somme du CUSUM au temps $t-1$ et à la différence entre la valeur observée et la valeur attendue au temps t . Le paramètre suivi par cette méthode est ainsi une somme d'écartés cumulés (Figure 4). Ci-dessous la formule de calcul du score CUSUM positif :

$$(1) C_0 = 0; C_t = \max(0; C_{t-1} + Y_t - k); t \geq 1$$

C_0 Cusum score à $t=0$

C_t Cusum score au temps t

Y_t Valeur observée au temps t

k Valeur attendue

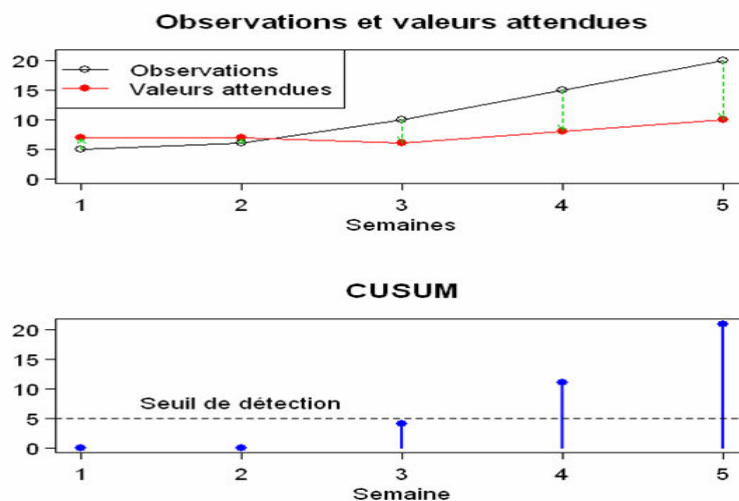


Figure 4 : Schéma du fonctionnement de la méthode CUSUM. La différence entre les valeurs observées et les valeurs attendues est calculée à chaque temps. La somme de ces différences depuis le temps 1, c'est-à-dire le CUSUM, est représentée par les barres bleues (Pelat 2010).

Deux périodes sont à définir par l'expérimentateur : la période de référence qui sera utilisée pour estimer la valeur attendue, et la période tampon qui est une période précédant l'observation et qu'on ne souhaite pas utiliser dans l'estimation de la valeur attendue. Une alarme est générée si la somme CUSUM dépasse une valeur seuil qui est définie par l'expérimentateur en fonction de la sensibilité et spécificité qu'il souhaite obtenir. La méthode permet de mettre en évidence des dérives graduelles dans le temps et pas seulement des modifications brusques.

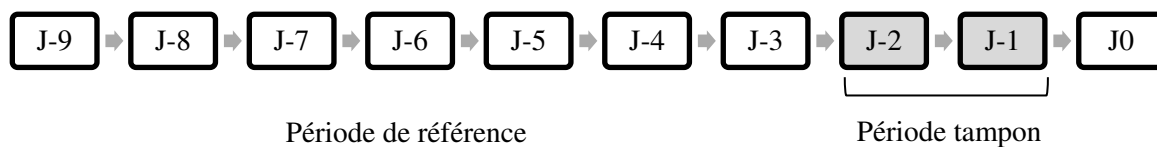


Figure 5: Illustration du choix d'une période de référence de 7 jours et d'une période tampon de 2 jours. L'observation courante est au temps 0.

Application de la méthode au suivi des ventes de médicaments

Un développement de la méthode CUSUM pour la surveillance des maladies infectieuses est l'utilisation d'un CUSUM basé sur une distribution négative binomiale. Watkins et al. a montré que le CUSUM binomial présentait des performances élevées pour la détection des épidémies au virus de Ross River (Watkins et al. 2008). On suppose que les cas suivent une distribution négative binomiale et les valeurs attendues sont calculées selon cette distribution. Nous avons appliqué la méthode sur les ventes d'anti-diarrhéiques et d'antiémétiques achetés sans ordonnance pour le traitement de la gastro-entérite. La figure 6 représente les quantités de ventes hebdomadaires entre 2007 et 2012. On souhaite détecter les signaux dans cette série temporelle, c'est-à-dire mettre en évidence les périodes épidémiques. On choisit comme période tampon 2 semaines afin d'éviter de prendre en compte un début d'épidémie dans l'estimation de la valeur attendue et une période de référence de 10 semaines. On s'intéresse uniquement aux changements positifs, la valeur minimale prise par le score CUSUM est 0. On choisit un seuil de détection à 0, qui est le scénario le plus simple. Chaque score de CUSUM supérieur à 0 met en évidence un signal anormal. Le programme a été développé sous R. Le score CUSUM calculé chaque semaine est représenté en rouge sur la Figure 6. La méthode de CUSUM permet de mettre en évidence huit périodes épidémiques correspondant aux pics estivaux et hivernaux dans les ventes de médicaments.

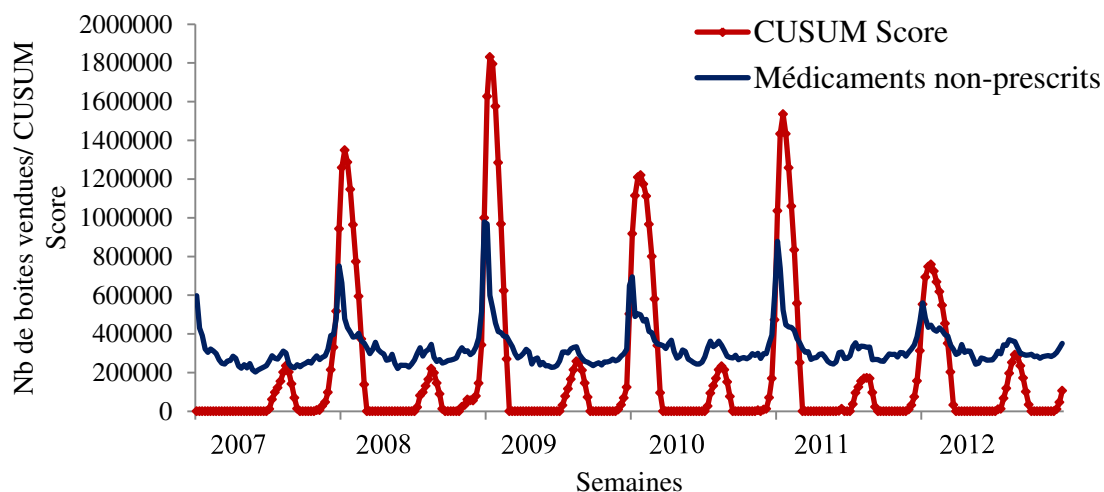


Figure 6: Nombre de boîtes de médicaments pour le traitement de la gastro-entérite vendues sans ordonnance par semaine (ligne bleue) et Score de Cusum calculé chaque semaine (ligne rouge). France Métropolitaine. 2007-2012.

Des périodes épidémiques sont détectées en été et en hiver avec la méthode CUSUM, alors que seules des périodes épidémiques hivernales sont mises en évidence avec les données du Réseau Sentinelles. Pour les pics hivernaux, la détection est en moyenne 4,5 semaines plus précoce avec la méthode CUSUM. La méthode permet de détecter de façon précoce les épidémies, cependant elle est très sensible et détecte de faibles augmentations.

(ii) La méthode EWMA

Principe de la méthode

La méthode EWMA (Exponentially Weighted Moving Average) utilise des données lissées, calculées à partir des moyennes mobiles pondérées exponentiellement (Roberts 1959) (Roberts 1959; Cisar et al. 2010). Le lissage permet de réduire le bruit dans la série temporelle et de limiter les faux positifs. Cette méthode donne également plus d'importance aux observations les plus récentes par rapport aux observations plus anciennes. Ainsi, à chaque temps t , la valeur lissée est calculée selon cette formule :

$$(2) \text{EWMA}_t = \lambda Y_t + (1 - \lambda) * \text{EWMA}_{t-1}$$

λ le paramètre de lissage. Il varie entre 0 et 1 selon le degré de lissage souhaité.

Y_t l'observation au temps t

EWMA_t la valeur lissée au temps t

La variance estimée de la statistique EWMA est :

$$\sigma^2_{EWMA} = \frac{\lambda}{2-\lambda} \sigma^2$$

σ écart-type calculé à partir des données historiques

On définit la limite supérieure de contrôle (UCL) :

$$(3) UCL = EWMA_0 + k \sigma_{EWMA}$$

$EWMA_0$ est la moyenne attendue estimée à partir des données historiques et k est un nombre positif définissant la hauteur de la limite de contrôle. Une observation est donc définie comme épidémique lorsque la valeur observée est supérieure à $\pm k$ écarts-types de la valeur attendue.

Application de la méthode au suivi des ventes de médicaments

Nous avons appliqué la méthode à des données de ventes hebdomadaires de médicaments non-prescrits indiqués pour la gastro-entérite entre 2007 et 2012 (médicaments sélectionnés dans notre article (Pivette et al. 2014b)). Pour définir le $EWMA_0$, nous avons calculé la moyenne des 30% de valeurs les plus basses, afin d'obtenir un niveau de vente en périodes « non-épidémiques ». Nous avons fait varier le paramètre de lissage λ entre 0,2 et 0,6. Le paramètre k a été testé entre 1 et 3. L'objectif était de définir si l'on pouvait détecter les épidémies de gastro-entérites avec une bonne sensibilité, spécificité et si la détection pouvait être précoce à partir de la méthode EWMA appliquée aux ventes de médicaments. Nous avons pris comme référence la semaine de détection définie par le réseau Sentinelle, avec une fenêtre de détection de 3 semaines avant et après. Nous avons calculé la sensibilité, le taux de fausses alertes et le délai de détection avec les différents paramètres (λ et k) (Tableau 1). La figure 7 représente les données de ventes lissées avec un coefficient de lissage de 0,3. Les trois seuils de détections avec $k = 1, 2, \text{ et } 3$ ont été représentés.

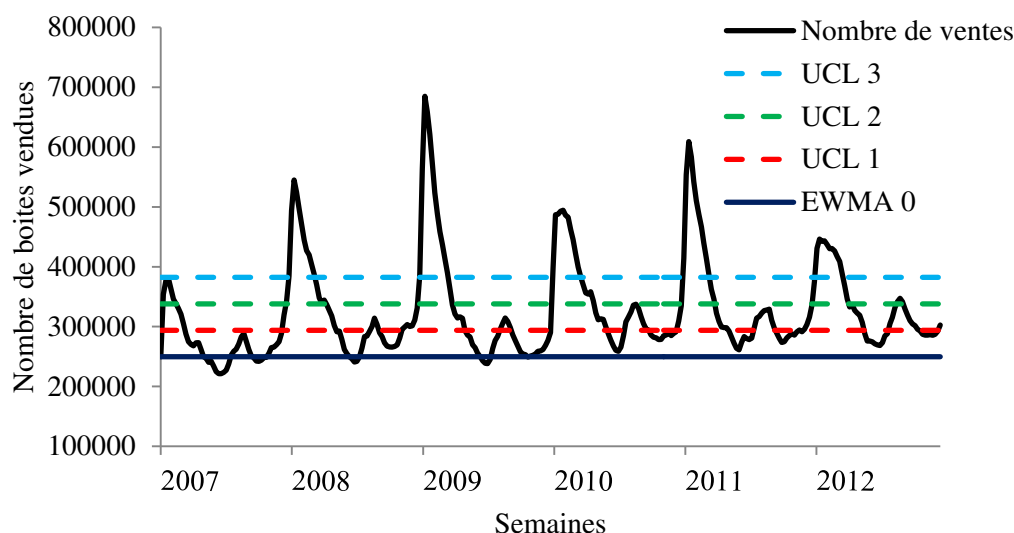


Figure 7 : Nombre de boîtes de médicaments sans ordonnance vendues pour le traitement de la gastro-entérite, par semaine entre 2007 et 2012, France Métropolitaine (données lissées avec un coefficient $\lambda = 0,3$). Les seuils de détection épidémiques définis à partir des coefficients $k=1$, $k=2$ et $k=3$ sont représentés par des lignes pointillées.

Tableau 1: Sensibilité, proportion de fausses alertes et délai de détection des épidémies de gastro-entérites à partir des ventes de médicaments non prescrits à partir de la méthode de détection EWMA.

λ	k	Sensibilité	proportion de fausses alertes	Délai (Semaines)
0,2	1	0,4	0,8	1,5
0,2	2	0,8	0,5	0,5
0,2	3	0,8	0,2	0,8
0,3	1	0,6	0,7	1,7
0,3	2	0,8	0,3	0,5
0,3	3	0,8	0,2	0,8
0,4	1	0,8	0,7	1,3
0,4	2	0,8	0,2	0,5
0,4	3	0,8	0,2	0,8
0,5	1	1	0,5	0,6
0,5	2	0,8	0,2	0,5
0,5	3	0,6	0,4	0
0,6	1	1	0,5	0,6
0,6	2	0,8	0,2	0,5
0,6	3	0,6	0,4	0

Les résultats des analyses montrent qu'aucune des combinaisons des paramètres (λ et k) ne permet de détecter toutes les épidémies de gastro-entérites et sans fausse alerte (Tableau 1).

(iii) Avantages et limites

Ces deux méthodes tiennent compte de l'historique des données et sont efficaces pour détecter de faibles changements dans une série de données. Elles peuvent être utilisées même si il n'existe qu'un court historique de données. L'avantage de la méthode CUSUM est de pouvoir détecter des petites déviations qui se succèdent dans le temps et pas seulement les fortes augmentations. Elle permet également une détection précoce du début d'une augmentation. Néanmoins, cette méthode est sensible à un faible changement du nombre de ventes et le risque de faux négatifs est élevé.

Ces méthodes de cartes de contrôle ont été conçues initialement dans le domaine du contrôle qualité dans la production industrielle afin d'émettre une alerte quand un procédé de fabrication sort des limites de contrôle. Elles sont donc indiquées lorsque la valeur attendue est constante au cours du temps et lorsque l'on s'attend à ce que l'observation montre une stabilité en l'absence de causes extérieures. Les deux méthodes testées avec les paramètres actuels empêchent de prendre en compte les tendances et les saisonnalités. Ainsi, une autre méthode de détection (modèle de régression périodique de Serfling) a été utilisée dans notre article (Pivette et al. 2014b) pour analyser les données de ventes de médicaments, en tenant compte de la tendance évolutive et des saisonnalités dans les données.

b) Méthodes basées sur les seuils historiques

(i) Méthode de Stroup

Cette méthode de détection a été utilisée par le CDC pour la surveillance des maladies à déclaration obligatoire (Stroup et al. 1989). Dans cette méthode, on considère comme valeur attendue pour une semaine s donnée d'une année a , la moyenne des valeurs observées dans les semaines $s-1$, s et $s+1$ des années précédentes. Un intervalle de confiance à 95% est calculé autour de cette moyenne et est utilisé comme seuil de détection.

Exemple d'une semaine comparée aux 15 valeurs des 5 années précédentes (Figure 8): Si la semaine d'intérêt est la semaine s_{15} en 2015 (valeur X_0). On calcule la moyenne \bar{X} des valeurs des cinq années passées pour la semaine s_{15} et les semaines adjacentes s_{14} et s_{16} . L'écart type attendu S est calculé. La borne supérieure de l'intervalle de confiance à 95% correspond au seuil de détection. La comparaison de la valeur observée X_0 à cette borne supérieure permet de définir si l'on est en période épidémique.

2015		X0	
2014	X1	X2	X3
2013	X4	X5	X6
2012	X7	X8	X9
2011	X10	X11	X12
2010	X13	X14	X15
	<i>s14</i>	<i>s15</i>	<i>s16</i>

$$\bar{X} = \sum Xi / 15$$

$$S = \sqrt{\frac{\sum (Xi - \bar{X})^2}{15}}$$

Seuil de détection (95th percentile)=
 $\bar{X} + 1,96 S$ (hypothèse normale)

$1,96 * \sqrt{\bar{X}}$ (hypothèse de Poisson)

Figure 8 : Illustration de la méthode de Stroup pour la semaine s15 de 2015 à partir des valeurs des 5 années précédentes.

Le nombre d'années et de semaines à prendre en compte pour le calcul de la valeur attendue dépend de l'application. Par exemple, le réseau Sentinelles utilise cette méthode pour calculer les taux attendus de grippe en été (Turbelin et al. 2009). Pour une semaine s, ils calculent la moyenne des incidences des années passées pour cette semaine s et les deux semaines autour (s-1 et s+1) depuis 1984 (Figure 9). Cette méthode est utilisée pendant les mois d'été par le Réseau Sentinelles au lieu de la méthode de régression de Serfling car l'intervalle de prédiction fixe ne permet pas de détecter les faibles augmentations en été (French GPs Sentinelles Network). Cette méthode permet de prendre en compte les saisonnalités dans les maladies infectieuses et a l'avantage de pouvoir détecter de faibles variations.

Cependant, cette méthode ne peut être utilisée que si les moyennes historiques utilisées pour calculer la valeur attendue ne présentent pas d'événements atypiques tels que les épidémies. En effet, la valeur attendue sera surestimée si des épidémies sont présentes lors des semaines utilisées dans le calcul de la valeur attendue. Il est donc nécessaire d'exclure les valeurs atypiques du calcul. Ainsi, cette méthode n'est pas adaptée pour la surveillance de la grippe ou la gastro-entérite en hiver qui présentent une période épidémique chaque année.

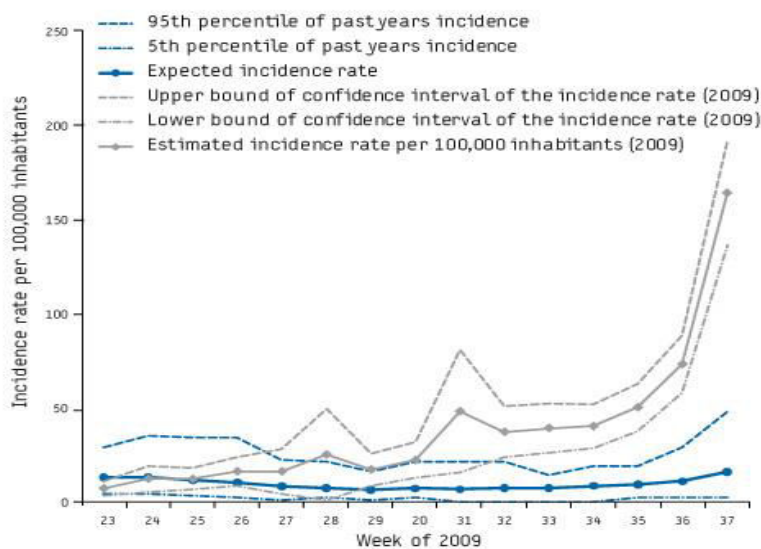


Figure 9 : Taux d'incidence hebdomadaire des syndromes grippaux en France (gris) et son intervalle de confiance, semaines 23 à 37 de 2009. Le taux d'incidence attendu (bleu) est calculé comme la moyenne des valeurs historiques (s-1 ; s ; s+1) des 24 années précédentes; le 95ème quantile des observations historiques (bleu pointillé) est le seuil de détection. Le seuil est dépassé lors de la semaine 31 (Turbelin et al. 2009).

(ii) Algorithme de Farrington

La méthode développée par Farrington et al. est basée également sur des données historiques mais une régression log-linéaire de Poisson est utilisée pour estimer la valeur attendue (Farrington and Beale 1993; Farrington and Beale 1996). Pour calculer le nombre de cas de Salmonelles attendus lors d'une semaine « s », Farrington et al. prennent en compte le nombre de cas de la semaine « s » et des 6 semaines adjacentes pour les cinq années précédentes (Farrington and Beale 1993). La valeur attendue est ensuite calculée en utilisant une régression log-linéaire sur les 35 semaines historiques, qui permet de prendre en compte la tendance, le cycle saisonnier et les épidémies passées. Dans cette méthode, un faible poids est attribué aux valeurs passées ayant des valeurs extrêmes, ce qui permet de limiter l'impact des épidémies passées dans l'estimation de la valeur attendue. Un intervalle de prédiction est calculé et la borne supérieure indique le seuil de détection. Cette méthode est utilisée depuis 1996 par le Communicable Disease Surveillance Center (CDSC) au Royaume-Uni pour la détection d'aberrations en routine dans les données hebdomadaires du nombre de micro-organismes rapportés.

4) Développement d'un indicateur de suivi de la gastro-entérite par *Celtipharm*

Suite aux résultats de recherche obtenus dans le cadre de la thèse, la société *Celtipharm* a mis au point un outil pour suivre en temps réel la propagation de maladies en France Métropolitaine: les Indicateurs Avancés Sanitaires (IAS®). Nous présenterons en particulier l'indicateur IAS® sur le suivi de la gastro-entérite qui fait suite à l'étude présentée précédemment (Pivette et al. 2014b).

a) Présentation des indicateurs IAS®

La société *Celtipharm* a développé des indicateurs de suivi de pathologies à partir des ventes quotidiennes issues de 4600 pharmacies en France Métropolitaine, soit 21% des pharmacies totales.

Huit syndromes ou indicateurs sont pour le moment suivis : la gastro-entérite, les syndromes grippaux, les allergies, les infestations par les poux, la gale, la couverture vaccinale chez les plus de 65 ans et la consommation de baclofène en France. Un indicateur sur le suivi de la vaccination Rougeole-Oreillons-Rubéole est en cours de développement. L'objectif de cet outil est de pouvoir suivre en temps réel la propagation de maladies ou d'évènements de santé à partir du suivi des ventes d'une sélection de médicaments. Des cartes sont présentées quotidiennement sur un site web à cette adresse <http://www.openhealth.fr/fr-fr/> depuis janvier 2014. Les données sont disponibles à J+1 et sont en open access.



b) Indicateur de suivi de la gastro-entérite

Un indicateur de suivi de la gastro-entérite a été développé (Auvigne 2014) à partir du suivi des 256 produits identifiés précédemment dans notre article (Pivette et al. 2014b) et indiqués en cas de gastro-entérites. Le suivi des ventes de médicaments est réalisé à partir du pourcentage du chiffre d'affaire lié aux médicaments sélectionnés sur l'ensemble du chiffre d'affaire des officines. L'ensemble des médicaments, prescrits et non-prescrits, ont été pris en compte. Un niveau de base est défini (base 100) comme le pourcentage de chiffre d'affaire à la fin du printemps (semaines 19 à 28), quand les gastro-entérites sont les moins fréquentes.

L'indice de base est calculé pour chaque Unité Géographique de Représentation (6311 hexagones de 10 km de côté en France Métropolitaine) et au niveau national. Un lissage géographique est utilisé pour représenter, pour une unité géographique donnée, une moyenne pondérée des valeurs observées alentour. Les pondérations sont décroissantes en fonction de la distance jusqu'à s'annuler à la distance de lissage de 100 km.

Un cartogramme par bassin de population est utilisé pour présenter les données (Figure 10). Le fond de carte classique est déformé proportionnellement à la densité démographique en se basant sur les données de population INSEE 2010. L'indice est à 100, ou un peu en dessous, pendant les périodes de l'année où les ventes de médicaments pour la gastro-entérite sont les plus faibles. Les valeurs de l'indice observées au pic de l'épidémie hivernale ont varié de 167 à 284 entre 2009 et 2014.

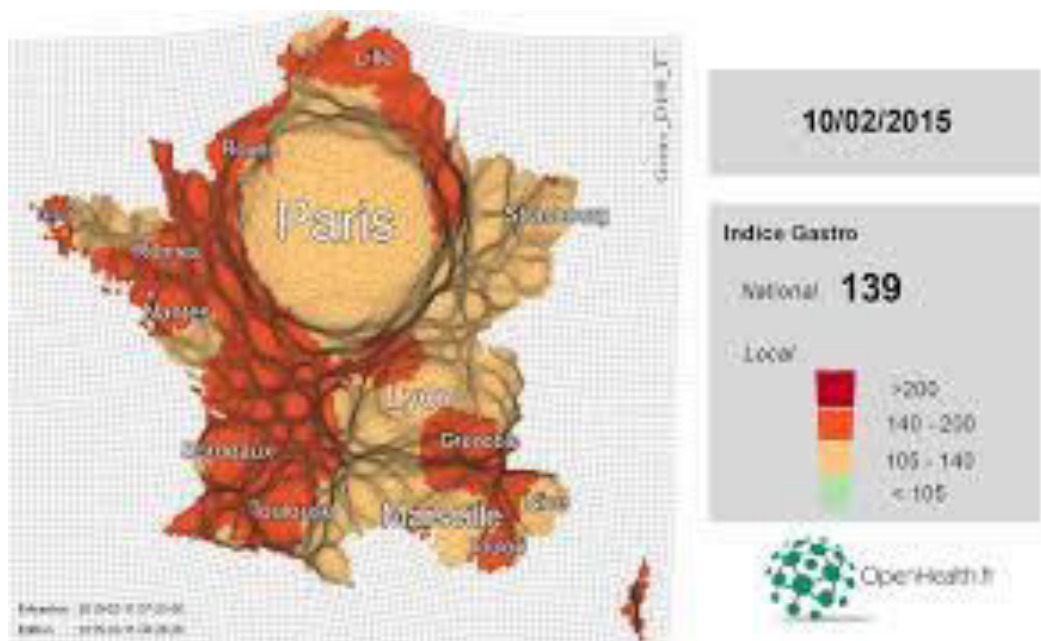


Figure 10: Représentation géographique de la gastro-entérite en France Métropolitaine (10/02/2015) à partir de l'indicateur développé par l'entreprise Celtipharm.

En plus des cartogrammes, les séries temporelles de l'indice sont également présentées au niveau national et au niveau de chaque région. Des bulletins hebdomadaires sont publiés sur le site www.openhealth.fr, et permettent à la société de communiquer sur les tendances observées des pathologies cibles.

c) Discussion

Ces travaux montrent qu'il est possible techniquement de publier à J+1 des données issues des

ventes en pharmacies et de fournir une information sur la consommation médicamenteuse en temps-réel. Les épidémies de gastro-entérites surviennent lors de chaque saison hivernale, cependant le début de l'épidémie est variable d'une année à l'autre. Il est donc intéressant d'avoir un outil permettant de visualiser la propagation de l'épidémie en temps réel. Le fait de présenter les résultats comme un pourcentage du chiffre d'affaire des médicaments indiqués pour la gastro-entérite sur le chiffre d'affaire total des pharmacies permet de pallier au problème de la détermination du dénominateur et permet d'avoir une indication sur l'évolution des ventes. Ainsi, les résultats donnent une indication sur l'évolution des ventes ; un indice supérieur à 100 indiquant une augmentation de la part du chiffre d'affaire liée aux ventes de médicaments contre la gastro-entérite par rapport au niveau de base (semaine 19 à 28). Les cartogrammes sont générés de façon automatique chaque jour et il est ainsi possible d'avoir une surveillance toute l'année et pas seulement pendant la période hivernale, comme cela a été fait dans notre étude.

Cependant, quelques limites sont à noter. Il n'est pas possible d'estimer une incidence de la maladie, les données proviennent d'un échantillon de 4600 pharmacies et ne sont pas extrapolables au niveau de chaque unité géographique de représentation. Dans ces modèles, il existe également un biais car il faut faire l'hypothèse que le niveau de base des ventes de produits contre la gastro-entérite dans chaque unité géographique se situe pendant la période des semaines 19 à 28. Une autre limite est que le dénominateur utilisé pour calculer l'indice est le chiffre d'affaire total des pharmacies, cependant celui-ci peut varier au cours de l'année en fonction des autres évènements de santé ou des épidémies. Ainsi, si une autre épidémie a lieu en même moment dans une région donnée, l'intensité de l'épidémie cible risque d'être sous-évaluée.

V. Estimation de la couverture vaccinale à partir des ventes de vaccins

Nous présenterons dans cette partie les résultats sur l'estimation de la couverture vaccinale contre la grippe et le rotavirus à partir des ventes de vaccins en France.

Les résultats sur la vaccination antigrippale des plus de 65 ans ont fait l'objet d'un premier article scientifique qui a été soumis à la *Revue d'Epidémiologie et de Santé Publique* en Janvier 2015 (M. Pivette, V. Auvigne, P. Guérin, J. Mueller, "Development and validation of a tool to estimate influenza vaccination coverage in real time in the population aged over 65 in France").

Les résultats sur la vaccination rotavirus sont des résultats préliminaires et les travaux sont en cours de développement.

1) Contexte et objectifs

Les recommandations vaccinales sont définies chaque année en France par le Haut Conseil de la Santé Publique et un calendrier des vaccinations est publié en fonction de l'âge et/ou des populations cibles (Ministère des Affaires sociales et de la Santé 2014). L'estimation de la couverture vaccinale (CV) et le suivi de l'activité vaccinale dans le temps sont nécessaires pour évaluer si les objectifs fixés par la loi de santé publique (DREES 2011a) sont atteints et si les programmes de vaccination sont correctement appliqués. Une couverture vaccinale suffisante permet l'élimination ou le contrôle de maladies infectieuses. Une CV insuffisante peut entraîner l'apparition d'épidémies et la résurgence de maladies, cela comme cela a été le cas en France lors de l'épidémie de rougeole en 2008-2011 (Antona et al. 2013). Le suivi de la couverture vaccinale est donc un élément essentiel pour guider les choix en matière de politique vaccinale et évaluer leur mise en application.

L'évaluation de la couverture vaccinale en France passe classiquement par l'analyse en routine des certificats de santé (0-2 ans), par des enquêtes de cycle triennal en milieu scolaire (2-15 ans) ou par des enquêtes occasionnelles réalisées dans des populations d'intérêt (Vaux et al. 2010; Guthmann et al. 2012). L'estimation est le plus souvent réalisée dans des populations spécifiques notamment chez les enfants et ne concernent pas la population générale. Les délais d'obtention des données sont importants et ne permettent pas un suivi réactif de la couverture vaccinale. Plus récemment, les données de remboursements issues de l'EGB ont été utilisées et permettent d'obtenir des données de couverture vaccinale de façon

plus réactive. Néanmoins, ces informations ne concernent que les vaccins remboursés par l'Assurance Maladie.

Notre objectif était d'évaluer l'intérêt du suivi des ventes de vaccins pour l'estimation de la couverture vaccinale et le suivi en temps réel de l'activité vaccinale en France. Les études existantes sur le suivi des ventes de vaccins en pharmacies ont principalement pour objectif de quantifier l'impact d'évènements tels que l'introduction d'un nouveau vaccin (Herzig et al. 1998; Simoens et al. 2009), la mise en place d'un changement de politique vaccinale (Guthmann et al. 2011) ou la mise en place de campagnes de sensibilisation (Mouly et al. 2013). Cependant, les données de ventes ne sont pas actuellement utilisées en routine pour estimer une couverture vaccinale.

Les données de vente de vaccins peuvent être obtenues rapidement (J+1) et donnent des informations sur l'ensemble de la population et tous les types de vaccins (prescrits et non-prescrits). Elles peuvent donc présenter un gain important en termes de réactivité et de représentativité de la population par rapport aux sources de données existantes.

Notre étude a porté sur deux vaccins : le vaccin antigrippal et le vaccin contre le rotavirus. Les données actuelles de la CV de la grippe proviennent de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAMTS) et sont calculées à partir des remboursements des bons de vaccination envoyés aux assurés du Régime Général ciblés par les recommandations qui sont les personnes de plus de 65 ans et les personnes souffrant d'affections de longue durée (ALD). Les données sont communiquées après la saison épidémique de grippe. L'objectif de notre étude était d'estimer la CV de toute la population et de pouvoir fournir une estimation en temps réel chez la population cible des plus de 65 ans lors de chaque saison épidémique.

La vaccination contre les rotavirus est indiquée chez les nourrissons de moins de 6 mois. Le vaccin rotavirus n'est pas remboursé et il n'existe actuellement pas d'estimation de la couverture vaccinale en France. L'objectif de notre étude était d'estimer la CV à rotavirus en France depuis 2008, et d'analyser les disparités géographiques de CV et leur association avec les hospitalisations pour gastro-entérites chez les enfants de moins de 2 ans.

2) Vaccination contre la grippe

a) Introduction

La grippe touche en moyenne 2,5 millions de personnes chaque année en France Métropolitaine (Invs 2015) et peut entraîner des complications graves et des décès chez les personnes à risque. L'épidémie saisonnière de grippe apparaît entre novembre et mars. Chaque année, une campagne de vaccination contre la grippe saisonnière est mise en place en France fin septembre/début octobre jusqu'à fin janvier. Le vaccin procure une immunité deux à trois semaines après l'injection. Pour être efficace, il est donc important que la vaccination soit réalisée au début de la campagne de vaccination. Une campagne de communication est mise en place à destination des professionnels de santé et du grand public (radio, presse, internet). L'Assurance maladie invite les assurés ciblés par les recommandations vaccinales du Haut Conseil de la Santé Publique à se faire vacciner via l'envoi d'un imprimé de prise en charge. Il s'agit notamment des assurés de plus de 65 ans, des personnes atteintes de certaines affections de longue durée (ALD), des personnes souffrant d'asthmes et de BPCO (Ministère des Affaires sociales et de la Santé 2014). Les médecins disposent également d'un imprimé de prise en charge qui leur permet de prescrire le vaccin aux personnes à risque n'ayant pas reçu de bon de vaccination telles que les patients obèses ou les femmes enceintes. Le vaccin est pris en charge à 100 % pour ces populations à risque munies d'un bon de vaccination pendant la période de campagne vaccinale. Le vaccin peut également être acheté sans prescription médicale en pharmacie mais n'est pas remboursé dans ce cas. Un objectif de 75% de couverture vaccinale (CV) des groupes à risque a été fixé en France par la loi de santé publique (DREES 2011b) mais cet objectif n'est pas atteint actuellement. La couverture vaccinale contre la grippe saisonnière des populations cibles était de 60,2% en 2009 et en baisse continue depuis, avec une couverture vaccinale de 48,9% en 2013 (Caisse Nationale d'Assurance Maladie 2014).

Les couvertures vaccinales des populations cibles actuellement disponibles proviennent de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAMTS) et sont calculées à partir des remboursements aux assurés du Régime Général (hors Sections Locales Mutualistes (SLM))(Caisse Nationale d'Assurance Maladie 2013; Caisse Nationale d'Assurance Maladie 2014). Les taux de vaccination sont communiqués après la campagne de vaccination. L'objectif de notre étude était d'estimer la couverture vaccinale dans la population générale depuis 2008, tous régimes confondus, et de valider une méthode pour estimer en temps-réel la

couverture vaccinale de la population cible des plus de 65 ans. L'obtention des couvertures vaccinales en quasi temps-réel peut permettre d'évaluer l'état de vaccination de cette population et d'ajuster les campagnes de communication.

b) Méthodes

Nous disposons des données extrapolées de ventes de vaccins contre la grippe en France Métropolitaine (Hors Corse). Les données sont obtenues à J+1 et permettent de distinguer les vaccins qui ont été prescrits et ceux qui ont été achetés sans ordonnance. La présentation du bon de vaccination en pharmacie par un patient entraîne la classification comme vente prescrite. L'âge du patient est connu pour les vaccins prescrits (ou achetés avec un bon de vaccination) et pour les vaccins non prescrits si le vaccin est acheté avec un médicament prescrit ou si l'âge du patient est enregistré dans la pharmacie. Les vaccins antigrippaux pris en charge dans le cadre des campagnes de vaccination chaque année et disponibles en pharmacie de ville ont été sélectionnés (Tableau 2). Il s'agit de vaccins à conditionnement unitaire. Une seule dose est nécessaire par personne (sauf pour les enfants de moins de 9 ans, pour qui 2 doses sont indiquées pour une première vaccination).

Tableau 2: Vaccins contre la grippe disponibles en pharmacie lors de chaque saison hivernale (Nom commercial et code CIP du produit)

Nom commercial	CIP	2008/ 2009	2009/ 2010	2010/ 2011	2011/ 2012	2012/ 2013	2013/ 2014
Aggripal®	3594738						
Fluarix®	3412973						
Immugrip ®	3338550						
Influvac®	3365216						
Vaxigrip®	3212990						
Mutagrip®	3201526						
Tetagrip®	3293617						
Gripguard®	3549382						
Previgrip®	3338722						

Pour la population générale, nous avons estimé la CV chaque saison entre 2008 et 2014 en prenant comme numérateur le nombre de vaccins vendus du 1^{er} septembre au 31 mars et au dénominateur la population en France métropolitaine le 1^{er} janvier de la saison épidémique (INSEE 2015). Nous avons déterminé la part de vaccins achetés sur prescription et achetés

sans ordonnance. Pour déterminer les comportements de vaccination, nous avons analysé le délai d'achat des vaccins lors de chaque saison, en prenant comme point de départ le début de la campagne de vaccination antigrippale.

Pour les plus de 65 ans, la couverture vaccinale a été définie comme le rapport entre le nombre de vaccins vendus aux 65 ans et plus entre le 1^{er} septembre et le 31 mars de chaque saison et la population des 65 ans et plus au 1^{er} janvier de la saison en cours. Afin de pouvoir obtenir une estimation de la couverture vaccinale en temps-réel, la population des plus de 65 ans au 1^{er} janvier de l'année n en cours a été calculée à partir de la population des plus de 64 ans de l'année n-1, à laquelle on applique des taux de mortalité et un solde migratoire estimés par l'INSEE (Chardon and Blanpain 2010). Les CV obtenues à partir des ventes de vaccins lors de chaque saison (Septembre-Mars) ont été comparées aux CV obtenues à partir des données de remboursements de la CNAMTS. Un taux de vaccination cumulé quotidien a été calculé lors de chaque saison, de 2009/2010 jusqu'à 2013/2014.

c) Population Générale

(i) Estimation de la couverture vaccinale

Les ventes hebdomadaires de vaccins ainsi que les périodes de campagne vaccinale entre 2008 et 2014 sont représentées sur la Figure 11. Dans la population totale, on observe une CV autour de 17% en 2008/2009 et 2009/2010, puis une baisse de la CV autour de 14% à partir de 2010/2011 (Tableau 3).

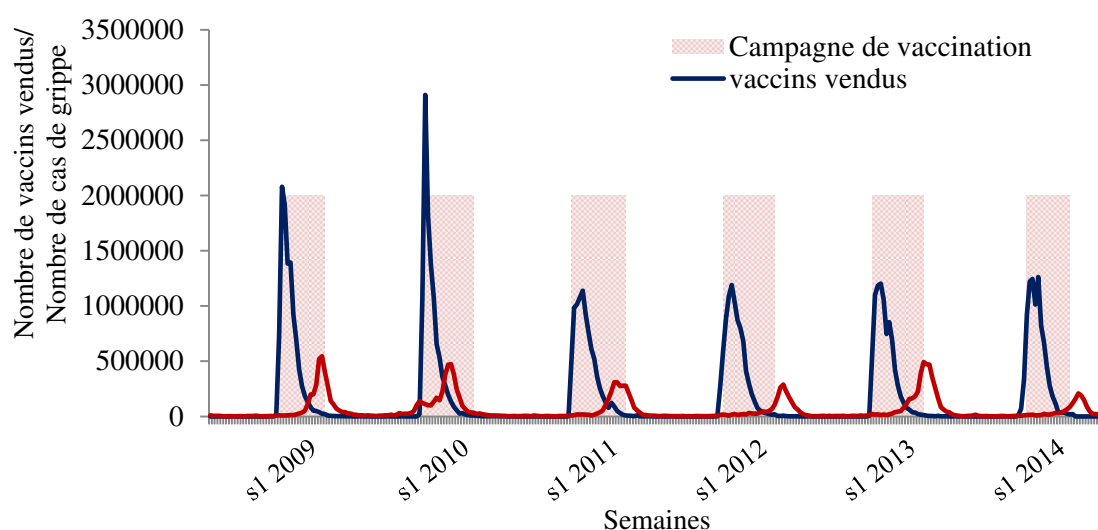


Figure 11 : Nombre de vaccins antigrippaux vendus en pharmacie, nombre de cas de grippe (Réseau Sentinelles) et périodes de campagne de vaccination. Semaine 17 2008 à semaine 35 2014, France Métropolitaine.

Tableau 3: Couverture vaccinale estimée lors de chaque saison hivernale (1^{er} septembre-31 mars). Population générale. France Métropolitaine.

	2008/ 2009	2009/ 2010	2010/ 2011	2011/ 2012	2012/ 2013	2013/ 2014
Couverture vaccinale (%)	16,7	17,1	14	13,8	13,6	13,7

(ii) Délai pour la vaccination

Les ventes cumulées journalières de vaccins pendant les campagnes de vaccination ont été analysées afin de déterminer le délai de vaccination (Figure 12). Pour étudier la cinétique de vaccination, le point de départ considéré était le premier jour de la campagne vaccinale chaque année (Tableau 4). Les vaccins prescrits et non prescrits ont été analysés séparément.

La part des vaccins achetés sans ordonnance représente en moyenne sur les 5 périodes de campagne vaccinale (2008-2013) 18,3 % des ventes de vaccins (19,1% en 2008 ; 22,9% en 2009 ; 17,3% en 2010 ; 16,4% en 2011 ; 15,7% en 2012). Le délai d'achat est en moyenne plus court pour les vaccins prescrits (75% des achats en 5,5 semaines en moyenne) par rapport aux non-prescrits (75% des achats en 6,9 semaines en moyenne).

On observe des disparités entre les années. Pour les vaccins prescrits, le délai est plus court avant 2010 (75% des achats en 4,5 semaines), qu'après 2010 (75% des achats en 6,2 semaines). Cette différence est aussi observée pour les non-prescrits, le délai est plus court avant 2010 (75% des ventes en 5,6 semaines) qu'après 2010 (75% des achats en 7,9 semaines). Environ deux semaines supplémentaires sont donc nécessaires pour atteindre 75% des achats de vaccins après 2010.

On remarque que le nombre de vaccins vendus est quasi nul chaque année à partir de début janvier. Il semble donc que prolonger la campagne, comme cela a été fait en 2010-2011 est une mesure d'efficacité faible.

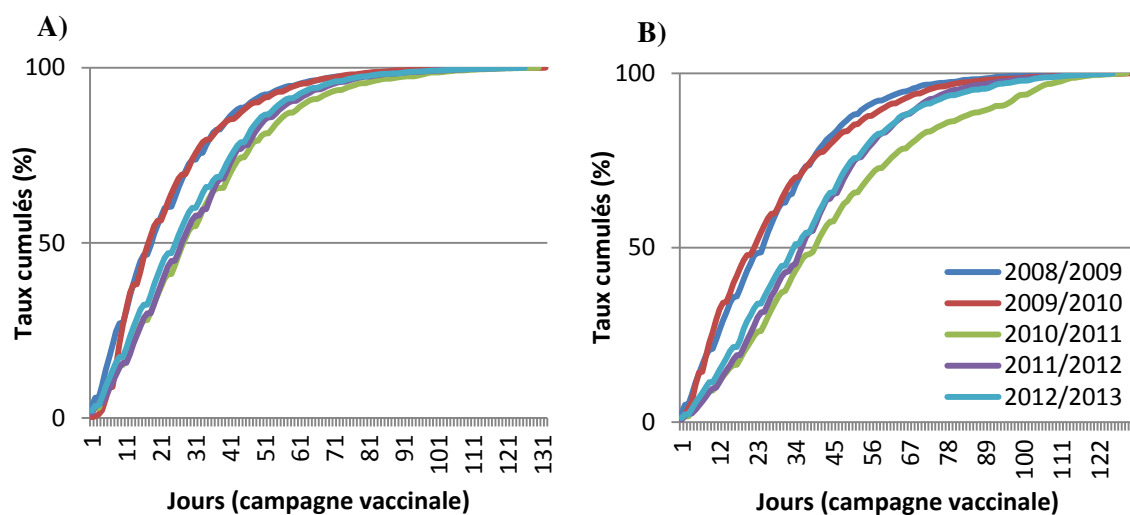


Figure 12: Taux cumulés journaliers de ventes de vaccins grippes lors de chaque campagne de vaccination. A) Vaccins prescrits B) Vaccins non-prescrits. Population générale, période de campagne vaccinale, 2008-2013, France métropolitaine

Tableau 4 : Dates de début et de fin de campagne vaccinale (2008-2013)

2008/2009	10 octobre - 31 janvier
2009/2010	23 septembre -15 février
2010/2011	18 septembre - 15 février
2011/2012	29 septembre - 31 janvier
2012/2013	28 septembre - 31 janvier

d) Population des plus de 65 ans

(i) Détermination et validation de la couverture vaccinale

Les plus de 65 ans sont une population ciblée par les recommandations et donc théoriquement qui achète le vaccin uniquement sur ordonnance (ou bon de vaccination) afin de pouvoir obtenir un remboursement. Nous pouvons donc estimer une couverture vaccinale pour les plus de 65 ans car l'âge du patient est enregistré pour les médicaments prescrits, cela n'est pas possible pour les autres populations à risque (patients en ALD, personnes obèses) qui ne peuvent pas être identifiées dans la base de données. Les plus de 65 ans représentent 78% des personnes invitées à se faire vacciner (Caisse Nationale d'Assurance Maladie 2014).

Les CV obtenues à partir des ventes de vaccins ont été comparées avec celles obtenues à partir des remboursements de la CNAMTS. Les CV obtenues à partir des remboursements sont définies comme le rapport entre le nombre de vaccins remboursés (datamart de consommation inter-régime (DCIR)) et le nombre de bons de prise en charge envoyés (Données ESOPE)

(Caisse Nationale d'Assurance Maladie 2013; Caisse Nationale d'Assurance Maladie 2014). Ces données concernent uniquement les personnes affiliées au Régime général hors sections locales mutualistes (SLM) en France Métropolitaine (Ministère des Affaires sociales et de la Santé 2013).

La comparaison a été effectuée sur 5 saisons (2009/2010 à 2013/2014). La CV estimée à partir des ventes de vaccins est en moyenne très légèrement plus élevée que celle estimée par la CNAMTS, pour chaque tranche d'âge (0,6 points supérieure en moyenne pour ≥ 65 ans ; 0,8 points pour les 65-69 ans ; 0,4 points pour ≥ 70 ans) (Tableau 5). La CV estimée en 2009/2010 chez les ≥ 65 ans à partir de ventes de vaccins était par exemple de 63,6% (6 685 012 vaccins vendus ; 10 508 701 personnes ≥ 65 ans). A partir des données de remboursements, la CV était de 63,9% (4 526 164 vaccinés/ 7 082 221 invités). Lors de la dernière année d'étude en 2013/2014, la CV chez les ≥ 65 ans à partir de ventes de vaccins était de 52,9% (6 118 064 vaccins vendus ; 11 574 034 personnes ≥ 65 ans). A partir des données de remboursements, la CV était de 51,9% (4 175 598 vaccinés/ 8 050 833 invités).

Une baisse de la CV est observée avec les deux sources de données entre 2009 et 2014 (-10,8% avec les données de ventes ; -12% avec les données de remboursement pour les ≥ 65 ans).

La corrélation entre les CV obtenues à partir des ventes de vaccins et celles obtenues à partir des remboursements est très forte lorsque l'on considère les plus de 65 ans (Spearman $r=1$) et les sous-groupes des 65-69 ans (Spearman $r=1$) et des plus de 70 ans (Spearman $r=1$).

Tableau 5 : Couvertures vaccinales (%) par groupe d'âge (plus de 65 ans), obtenues par les ventes de vaccins et les données de remboursements de la CNAM. 2009-2014, France Métropolitaine.

CV (%)	65-69 ans			≥ 70 ans			≥ 65 ans		
	Ventes	CNAM	Diff.*	Ventes	CNAM	Diff.*	Ventes	CNAM	Diff.*
2009/2010	55,1	55,3	-0,2	66,3	66,8	-0,5	63,6	63,9	-0,3
2010/2011	44,9	43,8	1,1	61,2	60,4	0,8	57,2	56,2	1,0
2011/2012	41,8	41,8	0,0	60,4	60,1	0,3	55,5	55,2	0,3
2012/2013	40,2	38,7	1,5	59,7	59,0	0,7	54,3	53,1	1,2
2013/2014	38,5	37,1	1,4	58,8	58,2	0,6	52,9	51,9	1,0
Diff. moyenne			0,8			0,4			0,6

* : différence entre les CV obtenus à partir des ventes de vaccins et des remboursements de vaccins

(ii) Suivi en temps réel de la vaccination chez les plus de 65 ans

Description de la cinétique des ventes

Pour les saisons 2010/2011 à 2013/2014, l'analyse du nombre de vaccins vendus chez les plus de 65 ans montre qu'il existe une phase ascendante dans le nombre de ventes pendant 2 à 4 semaines, puis on observe un plateau pendant 3 à 4 semaines avec un nombre de ventes hebdomadaires de vaccins vendus entre 800 000 et 900 000 (Figure 13). Le nombre de vaccins vendus diminue ensuite progressivement pour atteindre une valeur nulle en semaine 51 ou 52. La saison 2009/2010, correspondant à la pandémie grippale, présente une dynamique différente avec une augmentation brusque du nombre de vaccins vendus à partir de la semaine 38 pour atteindre un pic en semaine 40 autour de 2 000 000 de vaccins vendus.

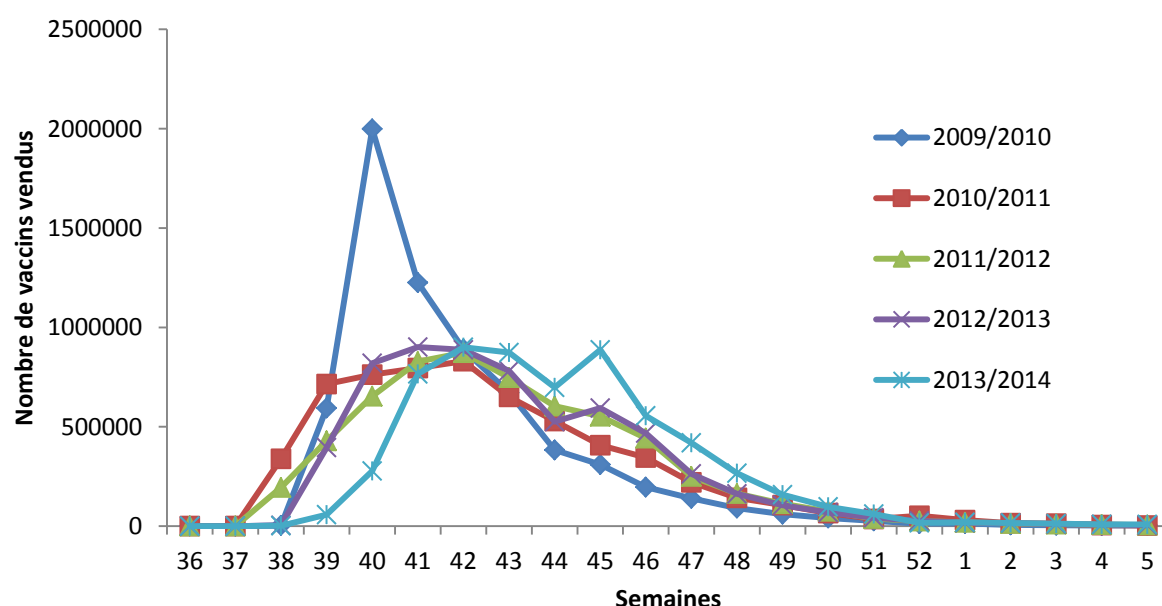


Figure 13 : Nombre de vaccins vendus par semaine entre la semaine 36 et la semaine 5, pour les saisons de 2009/2010 à 2013/2014. Population des ≥ 65 ans, France Métropolitaine (Hors Corse).

Suivi en temps réel de la couverture vaccinale

Les données de ventes sont collectées quotidiennement et disponibles à J+1. De plus, nous pouvons estimer la population des plus de 65 ans au 1^{er} janvier de l'année en cours en se basant sur la population de l'année précédente et de taux de mortalité et un solde migratoire estimés. Il est donc possible d'estimer la CV chez les plus de 65 ans chaque jour de la

campagne vaccinale. La couverture vaccinale quotidienne cumulée entre 2009 et 2014 a ainsi pu être estimée quotidiennement (Figure 14).

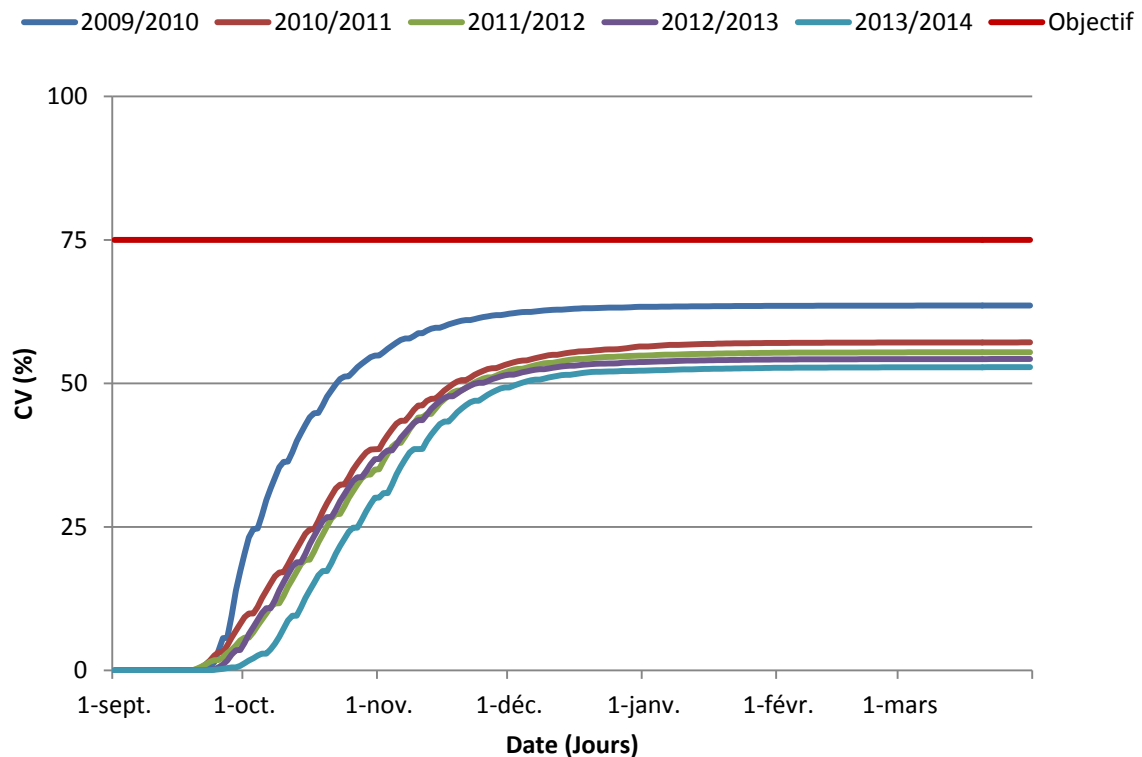


Figure 14: Couverture vaccinale cumulée journalière lors de chaque saison (1^{er} septembre- 30 mars) entre 2009/2010 et 2013/2014. Population des ≥ 65 ans, France Métropolitaine (Hors Corse).

(iii) Développement d'un indicateur par la société Celtipharm

A partir de ce travail, la société Celtipharm a développé un indicateur de suivi en temps-réel de la vaccination grippe chez les plus de 65 ans (<http://www.openhealth.fr/fr-fr/IASVaccGrippe65>) permettant d'accéder aux taux de vaccination cumulés quotidiens chez les plus de 65 ans en quasi temps réel, à J+1. L'utilisateur peut visualiser la couverture vaccinale quotidienne lors des années précédentes à partir de 2009/2010 et en temps réel pour l'année en cours. Les couvertures vaccinales journalières sont également accessibles en fichiers texte. Les données sont mises à jour quotidiennement. Les résultats sont également communiqués dans le bulletin hebdomadaire diffusé sur le site Openhealth (www.openhealth.fr) et la société Celtipharm a pu communiquer sur les taux de vaccination grippe chez les plus de 65 ans dans les médias pendant la saison épidémique 2014/2015.

Cet indicateur a été présenté dans l'article que nous avons soumis dans la Revue d'Epidémiologie et de Santé Publique (M. Pivette, V. Auvigne, P. Guérin, J. Mueller, "Development and validation of a tool to estimate influenza vaccination coverage in real time in the population aged over 65 in France")

e) Discussion

L'analyse des ventes de vaccins en officine permet d'estimer la couverture vaccinale de la population générale sans limite d'âge et d'identifier les tendances évolutives de la vaccination grippe. La couverture vaccinale a diminué après 2009/2010 (17% en 2009/10 et 14% en 2010/11). Cette baisse de la vaccination est probablement liée à la polémique suite à la délivrance du vaccin H1N1 en 2009/2010, comme cela a été montré dans des études antérieures sur les populations cibles (Tuppin et al. 2012a; Caille-Brillet et al. 2013). Cette baisse est maintenue dans les années suivantes. Les données de ventes permettent de cibler une population plus large que les données de remboursements existantes. En effet, toutes les ventes sont enregistrées, à la fois les vaccins prescrits et non prescrits et pour l'ensemble de la population et pas seulement les populations cibles.

On observe que les vaccins sont achetés plus rapidement lorsqu'ils sont prescrits que lorsqu'ils sont achetés sans ordonnance. Les personnes ciblées par les recommandations reçoivent un bon de vaccination au début de la campagne vaccinale, ce qui les incite à se faire vacciner rapidement et explique ce délai. On observe un retard dans l'achat au vaccin après 2009/2010. Les patients achètent les vaccins environ 2 semaines plus tard après 2009/2010, que ce soit pour les vaccins prescrits et les non-prescrits. L'analyse des délais est importante car il faut deux à trois semaines pour que la protection immunitaire induite par le vaccin soit efficace, une vaccination tôt permet une meilleure protection de la population. L'obtention des données quotidiennes en temps réel permet d'évaluer l'état de vaccination de la population et la dynamique de vaccination lors de chaque saison.

L'analyse des ventes de vaccins chez les patients de plus 65 ans, ciblés par les recommandations, permet d'estimer en temps réel la couverture vaccinale dans cette population. Les taux observés sont très corrélés à ceux obtenus à partir des remboursements de l'Assurance Maladie pour les affiliés au Régime Général hors SLM ayant reçu un bon de vaccination. Nos taux de CV sont légèrement plus élevés (0,6 point en moyenne). Cette différence peut s'expliquer car la population ciblée est différente, les données de

remboursements concernent les affiliés au Régime Général (hors SLM), qui représentent en moyenne 68% des plus de 65 ans (8 050 833 personnes au 1^{er} janvier 2014). Les données de ventes concernent l'ensemble des plus de 65 ans tous régimes confondus (11 574 043 personnes au 1^{er} janvier 2014). Le comportement de vaccination de personnes affiliées à d'autres régimes peut être différent de ceux affiliés au Régime Général. Cette hypothèse d'une couverture vaccinale supérieure dans les autres régimes doit être confirmée par l'étude des données de remboursements des autres régimes de l'assurance maladie (RSI, Régimes agricoles, Régimes spéciaux et Section SLM du RG). De plus, les données *Celtipharm* contiennent des achats effectués sans ordonnance et donc non-remboursés, alors que les données CNAMTS se basent uniquement sur les remboursements. On observe que les CV calculées en prenant en compte uniquement les vaccins prescrits sont en effet 1 point plus faibles que lorsque l'on étudie l'ensemble des vaccins vendus.

Actuellement, les données proviennent de la CNAMTS et sont communiquées après la saison grippale. L'intérêt d'une estimation journalière de la couverture vaccinale chez les plus de 65 ans est de pouvoir diffuser l'information en quasi temps-réel. L'indicateur développé par *Celtipharm* peut par exemple permettre de diffuser l'information de façon journalière au grand public, décideurs en Santé Publique, soignants et peut être utilisé pour évaluer et ajuster les messages de communication pendant la campagne de vaccination saisonnière.

Les données concernent uniquement les ventes effectuées en officines et n'incluent pas les vaccins distribués hors de l'officine de ville (par exemple les hôpitaux, médecine du travail, centres de vaccination, PMI, maisons de retraite avec Pharmacie à Usage Intérieur). Nous avons donc estimé une couverture vaccinale minimale. Cependant, les données basées sur les remboursements présentent cette même limite, car elles ne tiennent compte que de l'utilisation des bons de remboursements auprès d'un pharmacien (Tuppin et al. 2009; Tuppin et al. 2011). Les ventes de vaccins permettent d'identifier le groupe à risque des plus de 65 ans, l'âge du patient étant enregistré au moment de la vente. Cependant, il n'est pas possible d'estimer la couverture vaccinale dans les autres groupes à risque, notamment les patients en ALD et pour lesquels un suivi des données de remboursement est nécessaire.

f) Conclusion

L'analyse des ventes de vaccins grippe est un outil valide pour estimer la couverture vaccinale dans la population générale et dans la population cible des plus de 65 ans au niveau national.

De plus, la couverture vaccinale peut être estimée en quasi temps-réel. Cette estimation permet d'évaluer l'état de vaccination de la population et peut être utilisée pour ajuster les messages lors des campagnes de vaccination. Les analyses ont été réalisées au niveau national uniquement et des analyses complémentaires sont nécessaires pour évaluer la validité de ces estimations à un niveau géographique plus fin.

3) Vaccination rotavirus

a) Contexte

Les infections à rotavirus sont la principale cause de gastro-entérites aiguës de l'enfant (Gleizes et al. 2006; Parez et al. 2012). Pour la saison hivernale 2009/2010, les gastroentérites aiguës (GEA) à rotavirus ont entraîné 181 000 consultations, 31 000 passages aux urgences, 14265 hospitalisations et 6 décès chez les enfants de moins de 3 ans en France (HCSP 2013). Deux vaccins pour la prévention des infections à Rotavirus sont commercialisés en France: un vaccin monovalent Rotarix® et un pentavalent Rotateq®, disponibles respectivement depuis 2007 et 2006. Le Haut Conseil de la Santé Publique n'avait pas recommandé la vaccination systématique des nourrissons en 2006 ni en 2010 (HCSP 2010) mais avait recommandé l'amélioration de la prise en charge et des pratiques de réhydratation orale. Suite à une réévaluation par le HCSP de l'impact de la vaccination sur la morbidité et mortalité de la gastro-entérite dans les pays industrialisés, la vaccination à rotavirus en France est recommandée depuis février 2014 (HCSP 2013). La vaccination contre les rotavirus est indiquée chez les nourrissons de moins de 6 mois selon un schéma vaccinal à 2 doses (2 et 3 mois de vie) pour le Rotarix® et à 3 doses (2,3 et 4 mois de vie) pour Rotateq®. Les vaccins sont à prescription obligatoire (Liste 1) et non remboursables.

Les vaccins étant non remboursables, la CV ne peut pas être déterminée à partir des données de remboursements. L'objectif de notre étude était premièrement d'estimer la CV à rotavirus depuis 2008, au niveau national et régional. Une deuxième partie avait pour objectif d'évaluer l'association spatiale et temporelle entre la CV à rotavirus et le taux d'hospitalisation pour GEA chez les enfants de moins de 2 ans au niveau régional.

b) Estimation de la couverture vaccinale

(i) Méthodes

Le nombre de vaccins Rotarix® et Rotateq® vendus en officine chaque année entre 2008 et 2013 a été collecté, les données de ventes étant disponibles uniquement depuis le 1^{er} janvier 2008. Le nombre théorique d'enfants vaccinés par an a été calculé, en supposant une vaccination complète à 3 doses pour Rotateq® et à 2 doses pour Rotarix®. La couverture vaccinale a été estimée comme le rapport entre le nombre d'enfants vaccinés par an et le nombre de naissances par an. Les estimations de CV ont été réalisées au niveau national et régional.

(ii) Résultats

Au niveau national, la couverture vaccinale annuelle à rotavirus est comprise entre 6,9% et 9,5% entre 2008 et 2013 (Figure 15 A). La CV sur l'ensemble de la période 2008-2013 est de 7,8%.

L'ANSM avait recommandé en mars 2010 de ne plus utiliser le vaccin Rotarix suite à la présence de circovirus porcine dans des lots de vaccins Rotarix. En novembre 2010, l'Agence Européenne du Médicament a conclu à l'absence de risque de Santé Publique. Ces mesures de précaution ont entraîné la baisse des ventes de Rotarix en 2010 (Figure 15 B), expliquant la baisse de CV en 2010.

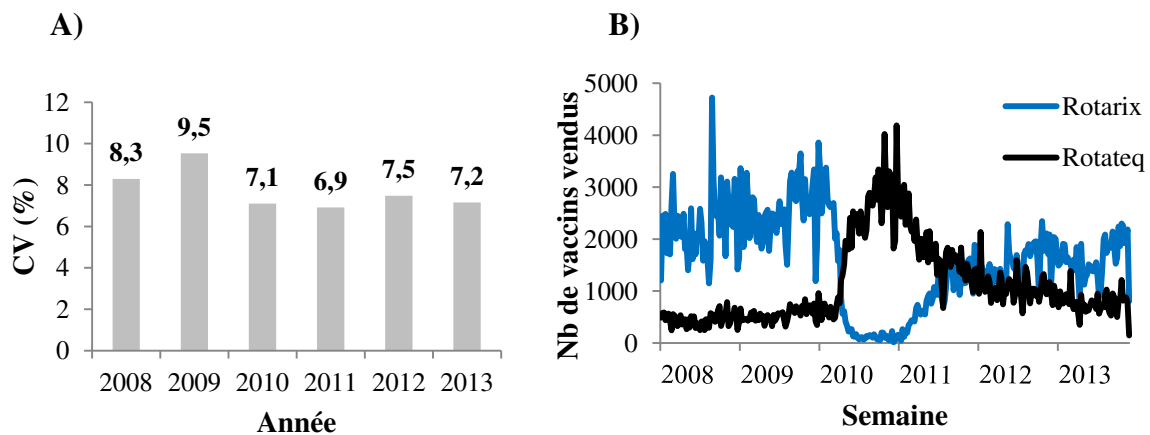


Figure 15: A) Couvertures vaccinales à rotavirus estimées par an entre 2008 et 2013. B) Nombre de vaccins vendus par semaine (Rotarix / Rotateq). France Métropolitaine (Hors Corse).

Au niveau régional, on observe de fortes disparités dans les couvertures vaccinales entre les régions (CV sur la période 2008-2012 Min : 4,1% ; Max : 12,4% ; Médiane : 6,6%). Les deux régions qui présentent le taux le plus fort sont l'Ile de France (12,4%) et l'Auvergne (12,2%) (Figure 16).

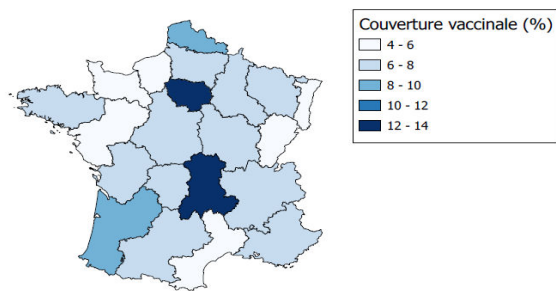


Figure 16: Couverture vaccinale à rotavirus par région (2008-2012). France Métropolitaine (hors Corse).

c) Hospitalisations pour gastroentérites à rotavirus

(i) Méthodes

Nous avons déterminé le nombre d'hospitalisations pour gastro-entérites aiguës à rotavirus chez les enfants de moins de 2 ans (0-23 mois) à partir des données du PMSI. L'infection à rotavirus présente un profil épidémique saisonnier et se propage pendant les mois les plus froids dans les pays tempérés (Cook et al. 1990; Parez et al. 2007). La période d'étude a donc été limitée de décembre à mai pour chaque saison épidémique (2008/2009 à 2011/2012).

Dans le PMSI, les maladies infectieuses et intestinales sont codées par des codes CIM-10 (Classification Internationale des Maladie) allant de A00 à A09. Le code « A08.0 : Entérites à rotavirus » existe mais la recherche du virus par un diagnostic microbiologique n'est pas effectuée systématiquement, et des gastro-entérites à RV peuvent être présentes sous d'autres codages, notamment sous le codage : Autres entérites virales (A08.3) et Infections intestinales virales sans précision (A08.4) (Fourquet et al. 2003).

Nous avons donc sélectionné dans notre étude les hospitalisations des enfants de moins de 2 ans ayant pour diagnostic principal un code CIM-10 : Entérites à Rotavirus (A08.0) ou Infections intestinales virales non précisées (A08.3, A08.4) afin d'augmenter la sensibilité. Les données PMSI étaient disponibles de 2008 à 2012.

Le taux d'hospitalisation a été calculé lors de chaque saison épidémique comme le rapport entre le nombre d'hospitalisations entre décembre et mai chez les enfants de moins de 2 ans et le nombre d'enfants de moins de 2 ans au 31 décembre de l'année (INSEE). Un taux annuel moyen d'hospitalisation a été calculé au niveau national et au niveau régional.

(ii) Résultats

Au niveau national, le taux global d'hospitalisations pour maladies infectieuses et intestinales (A00-A09) augmente légèrement (+ 11%) entre 2008 (1654 hosp/100 000) et 2012 (1836 hosp/100 000) (Figure 17). La part d'entérites à rotavirus (A08.0) parmi les maladies intestinales augmente (+22%) et le part d'infections virales non identifiées (A08.3 ; A08.4) diminue (-17%) (Figure 17). Une hypothèse à cette observation est l'augmentation des tests diagnostic à rotavirus pendant la période d'étude, qui a permis d'identifier des infections à rotavirus qui étaient classées comme infections virales non définies avant. Globalement, le taux d'hospitalisation pour entérites à rotavirus et infections virales non précisées augmente de 15% entre 2008 (822 hosp/100 000) et 2012 (950 hosp/100 000).

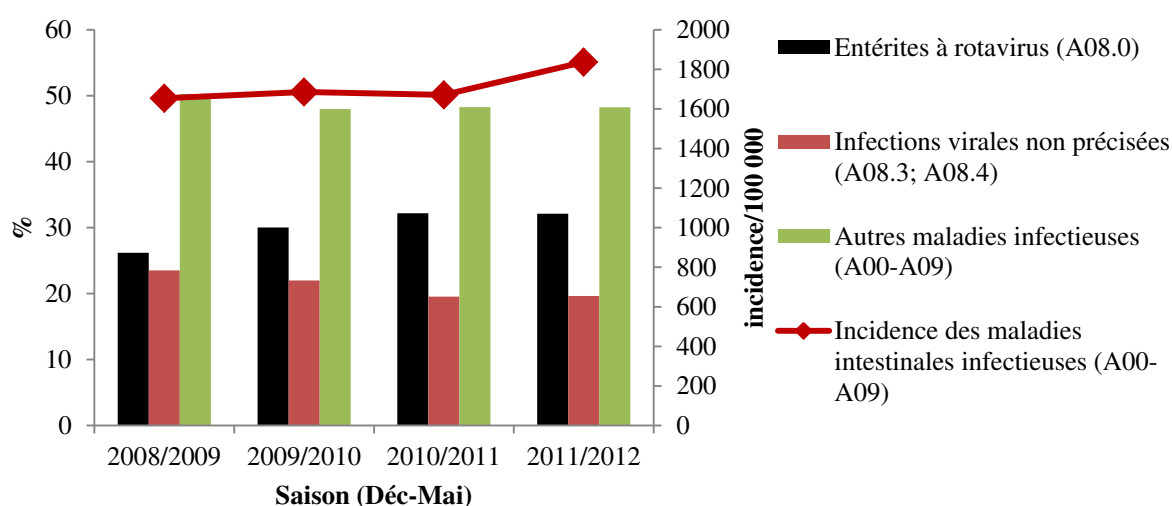


Figure 17: Proportion de chaque type de gastro-entérites (%) et incidence totale des hospitalisations pour Maladies infectieuses et intestinales (A00-A09) (ligne rouge) entre 2008 et 2012 chez les enfants de 0 à 2 ans. France Métropolitaine, décembre à mai pour chaque saison.

Au niveau régional, il existe de fortes disparités géographiques dans le nombre d'hospitalisations pour Entérites à rotavirus et Infections virales non précisées (min : 550; max : 1700 ; médiane : 933 hospitalisations /100000 enfants de 0-2 ans par an) (Figure 18). Les deux régions présentant des taux d'hospitalisation les plus élevés sont la Franche-Comté (1700 hosp/100000) et la Basse-Normandie (1539 hosp/100000).

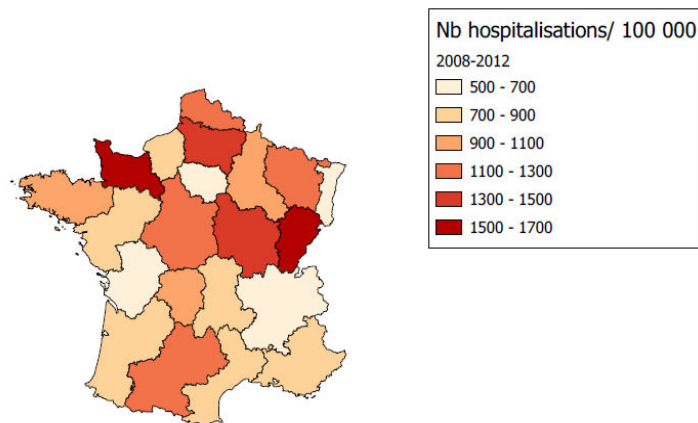


Figure 18: Taux annuel moyen (pour 100000) d'hospitalisations pour Entérites à Rotavirus et Infections intestinales virales non précisées chez les enfants de 0 à 2 ans, par région, France Métropolitaine (2008-2102).

d) Corrélation entre couverture vaccinale à rotavirus et les hospitalisations

Afin de quantifier l'impact de la vaccination sur le nombre d'hospitalisations pour gastro-entérites aiguës à rotavirus chez les enfants de moins de 2 ans, nous avons analysé la relation entre la couverture vaccinale et le taux d'hospitalisation pour Entérites à Rotavirus et Infections virales intestinales non précisées par région (Figure 19).

On observe que le taux annuel moyen d'hospitalisations pour rotavirus et Infections virales non précisées est plus faible lorsque la couverture vaccinale est élevée (en moyenne 633 hosp. /100 000 lorsque la CV > 12%) que lorsque la CV plus faible (en moyenne 1047 hosp. / 100 000 lorsque la CV < 8%). De plus, on observe que le taux annuel d'hospitalisations est plus dispersé lorsque le niveau de CV est plus faible (écart-type $\sigma = 304$ si CV < 8%) que lorsque le niveau de CV est élevé (écart-type $\sigma = 83$ si CV > 12%). Le taux d'hospitalisation et la CV étant connus pour chaque région de France, il n'y a pas de variabilité statistique liée à l'échantillonnage et nous n'avons pas à réaliser de test statistique pour comparer les deux groupes. Il faut noter que le nombre de régions où la couverture vaccinale est supérieure à 12% est très faible (n=2), ce qui limite l'interprétation de ces observations.

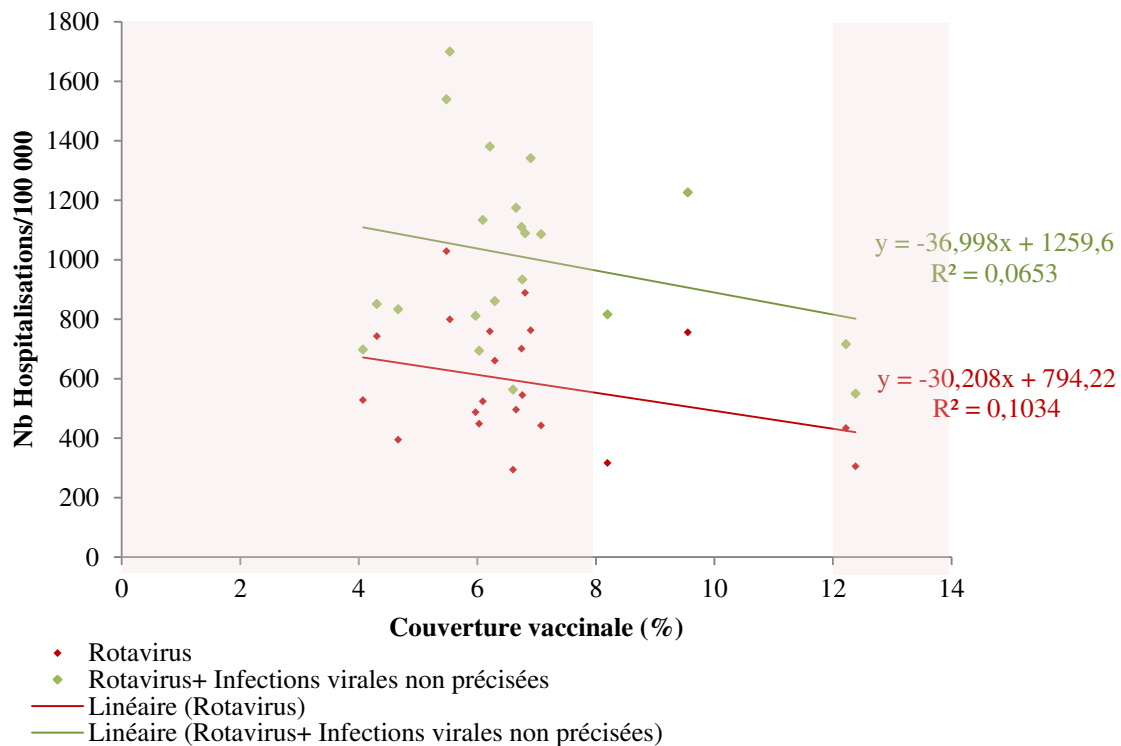


Figure 19 : Corrélation entre le taux annuel moyen d’hospitalisations (/100000) pour Entérites à Rotavirus et Infections intestinales virales non précisées chez les enfants de 0-2 ans et la couverture vaccinale moyenne à Rotavirus, par région, France Métropolitaine, 2008-2012. Chaque point représente une région.

Les équations des régressions linéaires simples réalisées entre les séries temporelles des taux d’hospitalisations et la couverture vaccinale sont indiquées sur la Figure 19. Lorsque l’on considère l’ensemble des hospitalisations (Rotavirus et Infections Virales non précisées), le coefficient de détermination est faible ($R^2 = 0,06$) et il n’y a pas de relation significative entre le taux d’hospitalisations et la CV ($p = 0,21$). Lorsque l’on considère uniquement les hospitalisations pour rotavirus, le coefficient de détermination est plus élevé ($R^2 = 0,10$) mais la relation n’est pas significative ($p = 0,06$) même si elle se rapproche de la significativité. On observe que la couverture vaccinale explique une plus grande partie de la variance lorsque que l’on étudie uniquement les hospitalisations pour rotavirus.

e) Discussion

L’analyse des données de ventes de vaccins Rotarix® et Rotateq® permet d’estimer une couverture vaccinale à rotavirus au niveau national et régional. Les vaccins étant non remboursés, il n’existe actuellement pas d’estimation de la couverture vaccinale et les

données de ventes permettent d'estimer cette CV autour de 7,8% sur la période 2008-2013, avec de fortes disparités géographiques. Le Haut Conseil de la Santé Publique n'avait pas recommandé la vaccination à rotavirus en 2006 et en 2010, mais avait recommandé l'amélioration de la prise en charge des pratiques de réhydratation orale. Cet avis peut expliquer la faible couverture vaccinale en France. De plus, le prix du vaccin (138 euros pour 2 doses de Rotarix®, 162 euros pour 3 doses de Rotateq®) peut être un frein à la vaccination. Suite à une réévaluation de la HCSP et l'observation d'une réduction de 80% des hospitalisations dans les pays industrialisés suite à l'introduction du vaccin (HCSP 2013), la vaccination à rotavirus en France est recommandée depuis février 2014. Le suivi des ventes de médicaments sera un outil utile pour évaluer en temps-réel l'impact de cette mesure sur la couverture vaccinale des nourrissons au niveau national et régional.

Les données d'hospitalisations étant disponibles uniquement depuis 2008, il n'est pas possible d'estimer l'impact de la mise en place de la vaccination sur le taux d'hospitalisations pour GEA à rotavirus. Une étude française (Gagneur et al. 2011), menée chez des enfants âgés de moins de 2 ans (1895 enfants vaccinés, 1985 non vaccinés) dans la ville de Brest, a montré que l'efficacité sur la réduction du risque relatif d'hospitalisation deux ans après le début de la mise en place d'un programme vaccinal, était de 98 % (IC95 : 84-100) après l'administration des trois doses de vaccin Rotateq®. Nous n'avons pas pu estimer la réduction en termes d'hospitalisations dans notre étude, cependant l'analyse des données de ventes de vaccins et des données d'hospitalisations pour gastroentérites à rotavirus au niveau régional permet de mettre en relation les deux séries de données. L'analyse des données montre que le nombre d'hospitalisations est très dispersé et élevé pour des CV faibles (CV < 8%). Cependant, au-dessus d'un certain niveau de CV (CV > 12%), le nombre d'hospitalisations est plus faible, ce qui est compatible avec l'existence d'un impact positif de la vaccination sur le taux d'hospitalisation. Cependant, le nombre de régions avec une CV élevé est faible (n=2), ce qui limite l'interprétation des résultats. De plus, aucune relation significative n'a pu être montrée entre la CV et l'incidence dans notre étude.

Notre étude présente certaines limites. En effet, nous avons considéré que la vaccination était complète, il s'agit donc de l'hypothèse basse de vaccination. De plus, la recherche de rotavirus n'étant pas systématiquement effectuée chez les patients hospitalisés, nous avons également analysé les infections virales non précisées, ce qui limite la spécificité de l'analyse.

f) Perspectives

Il s'agit d'une étude préliminaire et différents facteurs peuvent avoir un impact sur le taux d'hospitalisation et sur la couverture vaccinale. Ils devront être pris en compte dans la suite du projet pour analyser pour précisément le lien entre la couverture vaccinale et le nombre d'hospitalisations au niveau régional. En effet, l'incidence des gastro-entérites peut varier entre les régions, et pourra être prise en compte en analysant le nombre de consultations pour GEA dans la population générale à partir des données de consultations auprès des médecins généralistes ou des ventes de médicaments indiqués pour la gastro-entérite en pharmacie. Les vaccins Rotarix® et Rotateq® n'étant pour le moment pas pris en charge par la Sécurité Sociale et étant seulement remboursés par certaines mutuelles, l'achat du vaccin est influencé par le niveau socio-économique de la population (Esparza-Aguilar et al. 2014), qui devra ainsi être pris en compte dans des analyses complémentaires. Enfin, les solutés de réhydratation orale (SRO) (Van Cauteren et al. 2012) sont des produits remboursés par la Sécurité Sociale depuis 2003 et indiqués en première intention chez les nourrissons pour éviter les déshydratations et limiter l'évolution vers des formes aiguës de gastro-entérite nécessitant une hospitalisation. L'étude des ventes de SRO en pharmacies à partir de l'échantillon de pharmacies de *Celtipharm* pourra donner une indication sur les différences de prise en charge des gastro-entérites au niveau régional.

VI. Autres applications du suivi des ventes de médicaments

Le suivi des ventes de médicaments en pharmacies peut avoir des applications autres que la surveillance en routine de maladies infectieuses.

L'analyse des ventes de médicaments peut être utilisée de façon ponctuelle pour décrire les comportements de santé de la population et évaluer leur impact lors d'une crise de santé publique. En effet, les médicaments peuvent être achetés pour traiter des pathologies mais ils peuvent aussi refléter des comportements de prévention de la population, notamment les médicaments qui peuvent être achetés sans ordonnance (Miller et al. 2006). Deux études sur ces problématiques, réalisées en parallèle du projet de recherche de la thèse, seront présentées. Dans une première étude (Crépey et al. 2013), nous avons analysé la consommation de produits à base d'iode pendant la crise nucléaire de Fukushima en mars 2011 afin d'analyser les comportements de santé de la population. Dans une autre étude, nous avons analysé si la pandémie grippale de 2009/2010 et les mesures de santé publique mises en œuvre (utilisation des gels hydro-alcooliques) en France avaient pu avoir un impact sur une autre maladie infectieuse, la gastro-entérite. L'objectif de cette deuxième étude était d'évaluer l'effet de mesures de prévention sur la dynamique d'une maladie infectieuse, grâce à une étude écologique.

L'obtention de données d'achats de médicaments sans ordonnance peut également être utile pour compléter des études basées sur les remboursements de médicaments et pour décrire de façon plus précise la consommation médicamenteuse. Dans une troisième étude portant sur la consommation de psychotropes chez les enfants et adolescents en France en 2010, nous avons analysé les ventes d'anxiolytiques et d'hypnotiques sans ordonnance afin de compléter les analyses basées sur les remboursements de ces médicaments par l'Assurance Maladie.

1) Evaluation des mesures de prévention en France pendant la crise nucléaire de Fukushima.

Ce travail a fait l'objet d'une publication scientifique en mars 2013 dans la revue PLoSOne:

P. Crépey, M. Pivette, A. Bar-Hen, "Quantitative Assessment of Preventive Behaviors in France during the Fukushima Nuclear Crisis," *PLoS One*, vol. 8, no. 3, pp. 1–5, 2013.

a) Résumé

La crise nucléaire de Fukushima en mars 2011 au Japon a créé des inquiétudes au niveau mondial sur le risque de radiations nucléaires. En Europe, les autorités de santé ont informé les populations sur l'inutilité de la consommation de produits à base d'iode. Un manque de confiance dans les messages officiels s'est développé récemment. Pour étudier l'impact d'une crise de santé publique sur les comportements inappropriés de la population, nous avons analysé les ventes de produits pharmaceutiques à base d'iode avant, pendant et après la crise de Fukushima.

Notre étude a porté sur le suivi des ventes de 58 produits composés d'iode, achetés avec ou sans ordonnance. Les données viennent de l'échantillon de 3004 pharmacies développé par *Celtipharm*. La période d'étude allait de janvier 2010 à avril 2012. La méthode de détection CUSUM a été utilisée pour détecter les augmentations dans les ventes et la méthode de cross-corrélation pour évaluer les tendances évolutives de chaque type de produits les uns par rapport aux autres.

Les ventes de compléments alimentaires à base d'iode et ensuite de produits homéopathiques à base d'iode ont augmenté (respectivement 1,5 fois et 3 fois plus élevés en Mars-Avril 2011 par rapport à la même période en 2010). Les pics observés dans le nombre de produits vendus sont reliés à des événements particuliers pendant la crise de Fukushima, notamment aux dates auxquelles les médias ont communiqué sur l'arrivée du nuage radioactif en France. Les médicaments prescrits homéopathiques à base d'iode ont également augmenté (35% d'augmentation).

Les résultats ont mis en évidence des comportements préventifs inadaptés qui ont abouti à la consommation injustifiée de médicaments. Cette étude montre l'intérêt du suivi des ventes de médicaments pour évaluer rapidement les comportements de la population pendant une crise globale. L'analyse de ces données peut être utile pour ajuster les messages de communication afin de limiter les comportements inadaptés de santé.

b) Article

Quantitative Assessment of Preventive Behaviors in France during the Fukushima Nuclear Crisis

Pascal Crépey^{1,2*}, Mathilde Pivette^{1,3}, Avner Bar-Hen³

1 EHESP Rennes, Sorbonne Paris Cité, Paris, France, **2** Aix-Marseille Univ, UMR EPV Emergence des Pathologies Virales –190, Marseille, France, **3** Université René Descartes-Paris5, MAP5, Paris, France

Abstract

Background: The Fukushima nuclear disaster has generated worldwide concern on the risk of exposure to nuclear radiations. In Europe, health authorities had to issue statements about the lack of usefulness of iodine based preventive treatments within their borders. However a lack of confidence in official messages has developed in various European countries due to recent perceived failures in managing public health crises. The lay population preventive behaviors in this context are largely unknown. Consequently, to examine the effects of public health crisis on lay behaviors leading to pharmaceuticals purchases, we studied the sales of iodine-based products in France before, during and after the crisis.

Methods: We focused our study on 58 iodine-based drugs available with and without a physician prescription. Our data came from a stratified sample of 3004 pharmacies in metropolitan France. Our study period was from January 2010 to April 2012, with a focus on March-April 2011. We differentiated sales of drugs prescribed by physicians from sales of drugs obtained without a prescription. We used a CUSUM method to detect abnormal increases in sales activity and cross-correlations to assess shifts in sales timing.

Results: Sales of iodine-based nutritional complements, and later sales of iodine-based homeopathic remedies, substantially increased (up to 3-fold) during a period of 20 days. Their temporal patterns were correlated to specific events during the crisis. Prescriptions for iodine-based homeopathy increased (up to 35% of all sales). Iodine pills, strictly regulated by health authorities, have also been sold but on a very small scale.

Conclusion: These results indicate uncontrolled preventive behaviors resulting in the potentially unjustifiable consumption of available drugs. They have implications in public policy, and demonstrate the usefulness of drug sales surveillance for instantaneous evaluation of population behavior during a global crisis.

Citation: Crépey P, Pivette M, Bar-Hen A (2013) Quantitative Assessment of Preventive Behaviors in France during the Fukushima Nuclear Crisis. PLoS ONE 8(3): e58385. doi:10.1371/journal.pone.0058385

Editor: Suminori Akiba, Kagoshima University Graduate School of Medical and Dental Sciences, Japan

Received: September 20, 2012; **Accepted:** February 4, 2013; **Published:** March 7, 2013

Copyright: © 2013 Crépey et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Pascal.Crepey@ehesp.fr

Introduction

On March 11, 2011, a magnitude 9 earthquake followed by a 15-m tsunami wave damaged the Fukushima Daiichi Nuclear Power Plant (NPP) in Japan [1]. The disaster provoked the meltdown of three reactors leading to a level-7 nuclear incident [2]. In the aftermath, Japan had to face the worst nuclear disaster since the destruction of Chernobyl's reactor number four in 1986. Nuclear disasters have global consequences, and populations are now highly aware of the potential public health consequences of radioactive discharges [3,4].

In France, a severe loss of confidence in political authorities has been observed in matters concerning nuclear risk following the Chernobyl event. A study on a sample of the general population in France by Carlé and colleagues [4] showed that 75% of the respondents believed they had not been told the truth about the Chernobyl fallout. The same study showed that neither government representatives nor journalists were considered by a large majority of respondents to be technically competent and likely to tell the truth about nuclear issues. In addition, 68% of the

respondents had no confidence in the abilities of the authorities to protect them in case of a major nuclear disaster.

This lack of confidence in health authorities was also observed during the last influenza pandemic. However, the reluctance to comply with official recommendations seems to be more related to differences in risk perception [5,6], which resulted in a lower than usual pandemic influenza vaccination uptake [7].

During the Fukushima crisis, contrary to multiple unofficial and nonmedical sources, all French official communications indicated that using iodine was not useful as a preventive measure in France [8]. Potassium iodide (KI) pills are only used in case of risk of direct exposure to radioactive iodine isotopes to saturate the thyroid glands with nonradioactive iodine. In France, KI pills are army-regulated and only available in military controlled pharmacies or in pharmacies located in a 10-km radius from one of the 19 French NPPs [9]. However, several drugs, such as antiseptics, nutritional supplements and homeopathic remedies, containing relatively low concentrations of iodine, can be purchased in pharmacies with or without prescriptions.

Our study focuses on sales of the previously described drugs, on their timing, and on correlations with the timing of events that may have impacted those sales. Implications in terms of public health campaigns are discussed.

Materials and Methods

We studied the changes in pharmaceutical sales of iodine-based drugs in France through a stratified random sample of 3004 pharmacies set up by the company Celtipharm [10]. Through their participation in the Xpr-So© system, the pharmacies have automatically reported their sales several times a day since 2007. Thanks to a regularly updated exhaustive database of the 22,458 active French pharmacies, a stratification to improve representativeness is performed on sales revenue (6 levels for global revenue and 4 levels per type of sales: prescribed drugs, OTC, and other type of sales), localization (5 geographic areas) and sales area (5 types, from rural to densely urban). Each stratum has a minimum of 30 pharmacies or is merged with neighboring strata. Sampling rates per strata are computed with the Neyman optimal allocation algorithm [11]. Extrapolations from the sample have been validated with data from drug manufacturers who distribute their products directly (and only) to pharmacies. We study sales data from January 2010 to April 2012 in order to include a reference period (the year 2010) and the period surrounding the Fukushima crisis. We compiled a list of 92 products containing or claiming to contain iodine based on a search for the pattern “iod*” in their name or principal ingredients. From that list, a total of 58 had been sold at least once in pharmacies from our sample during the study periods. We analyzed separately sales made with a physician prescription and those made without prescriptions. Sales were observed at the national and regional levels. The number of pharmacies selling the listed products was also analyzed. The list of products is given in Table S1. We computed cross-correlations of sales time series to quantify potential lags between types of sales. To detect unexpected increases in sales and assess durations of this kind of event we used a negative binomial CUSUM algorithm. We parameterized the CUSUM to compute its score from the difference between the cumulative sum of the sequence of observations and the maximum of variations observed the preceding month. We used a lag of a few days between the current observation and the preceding month to avoid tainting the “variation baseline” with the observed event. The same method, developed for real-time detection of a parameter’s deviation, is also classically used to detect disease outbreaks [12,13]. Positive CUSUM score can occur during a normal year; consequently, to maximize the specificity of our indicator we set the detection threshold to the maximum value of the score obtained during an *a priori* “normal” year, i.e. the year 2010, preceding the Fukushima crisis.

Results

We defined four categories of iodine-based products: antiseptics, nutritional complements, homeopathic remedies, and KI pills. Figure 1 shows increases in sales in all product categories except iodine-based antiseptics during the period of the Fukushima crisis (February to mid-April 2011). In addition, although KI pills were only available to people living in a 10 km-radius from a NPP and were not released to the rest of the population for preventive purpose during the period, sales were registered in our sample specifically during the crisis. To confirm these increases, we computed a negative binomial CUSUM score [12] on the sales time series (Figure 2). We define the threshold for “sales outbreak” detection as the maximum CUSUM score obtained for the year

2010. Antiseptics (prescribed or not) do not cross the threshold in 2011, while iodine-based homeopathic remedies cross it during 22 days for all sales and 7 days for prescribed sales. Lastly, nutritional complements cross it during 19 days for all type of sales. Prescribed nutritional complements also slightly cross the threshold in 2011 but a few days prior to the Fukushima crisis. Sales of KI pills were almost nil outside of the Fukushima period making the CUSUM computation irrelevant. Comparisons between sales in 2010 and 2011 are given in Table 1. In particular, we notice a 144% increase of iodine-based nutritional complement sales in March 2011, but the largest sales increase concerns iodine-based homeopathic remedies, with a threefold increase compared to March 2010. We assessed the scale of the phenomenon by analyzing the number of pharmacies per day selling the identified products. In 2010, on a given day only 210 pharmacies on average sold iodine-based homeopathic remedies. During the 1-month period starting on March 11th, this number rose to 740 selling pharmacies per day on average and to 2,046 between the 21st and

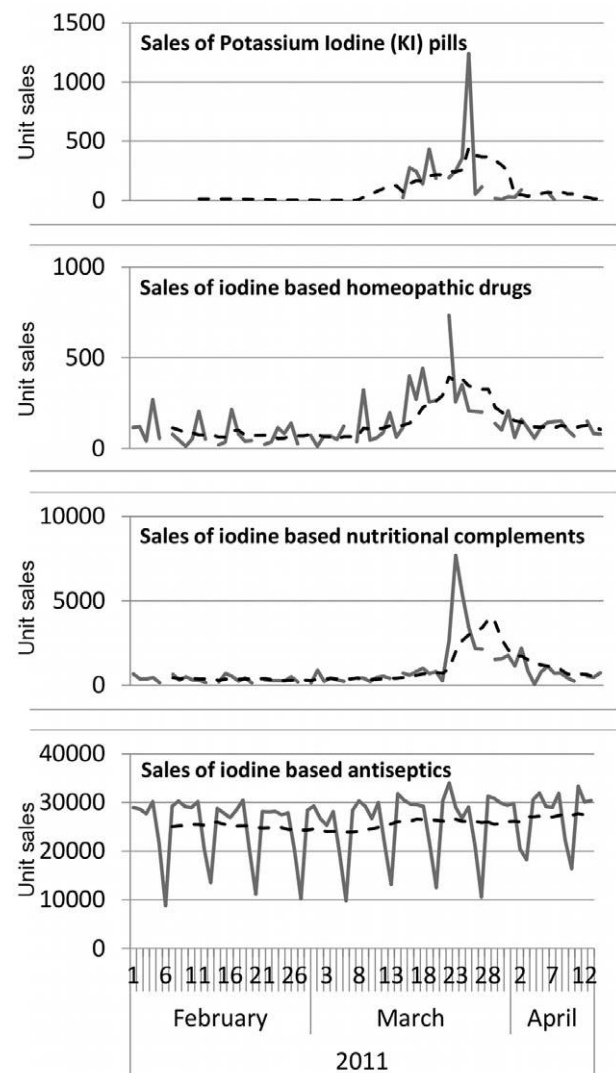


Figure 1. Sales of iodine based products (per units) in French pharmacies between February and April 2011. Dashed lines show the weekly sliding average.

doi:10.1371/journal.pone.0058385.g001

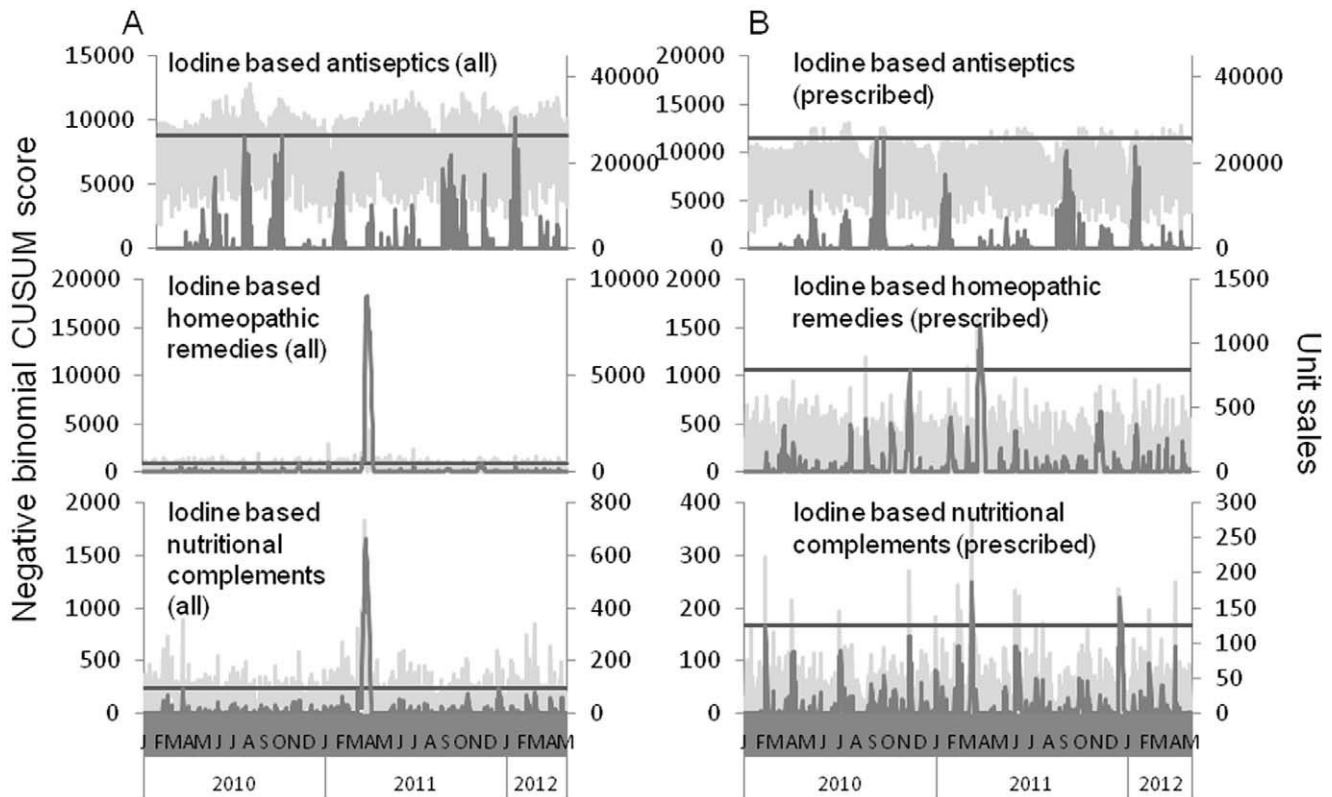


Figure 2. Dark gray lines stand for negative binomial CUSUM score for unit sales of iodine based antiseptics, homeopathic remedies and nutritional complements (from top to bottom). The straight lines on the plots illustrate the threshold for the CUSUM score, defined as the maximum value of the CUSUM score in 2010. An increase in the CUSUM score indicates a strong positive deviation from the normal trend. Light gray lines represent unit sales volumes. Panel A shows results for all type of sales while panel B only shows sales with a physician prescription.

doi:10.1371/journal.pone.0058385.g002

25th of March 2011, representing 10% of the total number of stores.

To explore the timing of the purchases, we computed cross-correlations between the different sales time-series displaying an increase between February and April 2011 (all categories except iodine-based antiseptics). The cross-correlograms (Fig. 3) show that nutritional complements were purchased first, then homeopathic remedies sales peaked one day later, followed two days later by a peak of KI pills sales. No regional patterns were evidenced in the data.

Discussion

The Fukushima crisis was a good model to study the public's reaction to a major public health crisis. Whether the risk was real

or not, we show that a part of the population chose to ignore official recommendations, seeking protection with what they perceived as preventive drugs. Table 2 allows us to hypothesize temporal links between these peaks and specific events related to the Fukushima crisis. The two main peaks of nutritional complement and homeopathic remedies sales occurred on the 18th and 24th of March, respectively, when a large-scale radioactive discharge was confirmed and just prior to the arrival of the radioactive cloud over France. The sequence of events, causing increasing concern, could explain the trend of purchase of different categories of products, which could be perceived of incremental efficiency: first iodine-based nutritional complements, then iodine-based homeopathic remedies, and finally KI pills.

Adverse events related to iodine consumption are limited since available studies do not reveal an association between KI and

Table 1. Comparison of iodine based products in one month in 2010 and the same month in 2011 after the beginning of the Fukushima crisis.

Estimated sales in French pharmacies (in units)	From March 11 to April 11 2010	From March 11 to April 11 2011	Percentage of increase
iodine based nutritional complement	2545	6219	144%
iodine based homeopathic remedies	10747	43912	309%
iodine based antiseptic	758436	825963	9%

doi:10.1371/journal.pone.0058385.t001

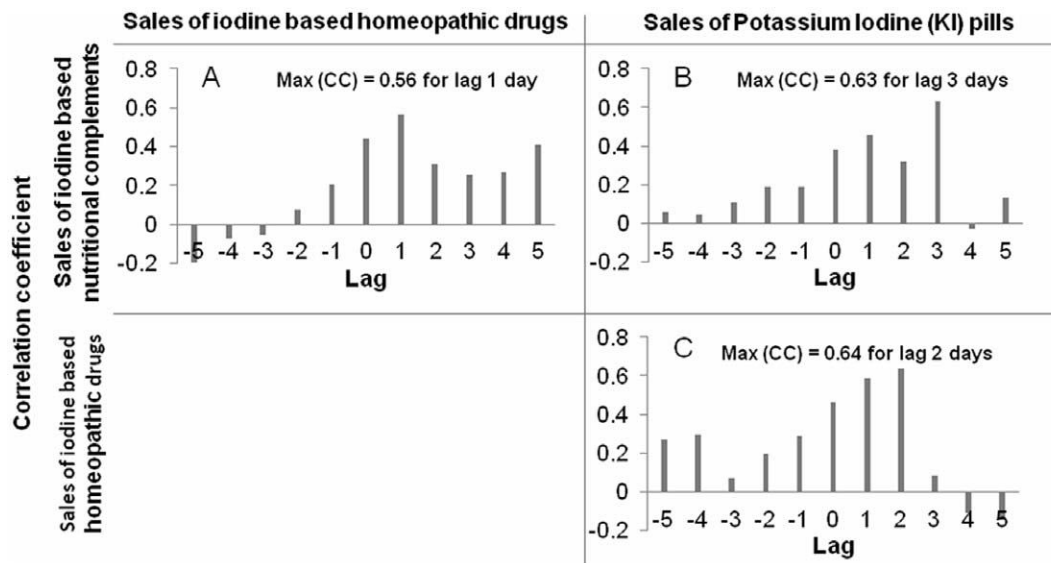


Figure 3. Cross-correlograms of sales of iodine based nutritional complements (INC), homeopathic drugs (H), and Potassium Iodide (KI) pills with lags from -5 to +5 days on the period from March 1st 2011 to April 10th 2011. a) Cross-correlations show that sales of INC preceded sales of H. by 1 day. b) Cross-correlations show that sales of INC preceded sales of KI by 3 days. c) Cross-correlations show that sales of H. preceded sales of KI by 2 days.

doi:10.1371/journal.pone.0058385.g003

severe adverse events in the general population. Only newborns, the elderly and people with a history of thyroid dysfunction are at higher risk of developing complications [14]. However our results raise questions about behaviors that could be adopted in the context of a public health crisis. The recent influenza A/H1N1 pandemic crisis in France showed that a part of the population chose to avoid protective measures. We now show that during the Fukushima crisis a part of the population looked for its own protective measures. It is also interesting to note that, despite the absence of any official indication or biological rationale, physicians have prescribed homeopathic remedies to protect their patients'

thyroidal glands. The generally admitted absence of adverse events on account of homeopathic drugs may have played a role in physicians' decision to prescribe them to demanding patients. This absence of adverse events could explain the rationale for such protective behavior, at no perceived cost in terms of health risk.

Our approach has some limitations: (i) Pharmacy sales are only a proxy for self-medication since we cannot quantify practices such as traditional medicines, internet drug sales, etc... (ii) The act of purchasing a drug does not imply that it has been consumed. It is possible that purchased drugs have been stockpiled in order to be available in case of a real threat. The strict regulation surrounding

Table 2. Timeline of events of the Fukushima crisis in Japan and in France from references [1,2,8].

Date	Japan	France
March 11 th , 2011	Magnitude 9 earthquake and tsunami damage the cooling systems of reactor 1, 2, 3 of the nuclear power plant. Evacuation order to persons within 3 km.	
March 12 th , 2011	Hydrogen explosion in unit 1 building. Evacuation order to persons within 10 km, then extended to 20 km.	
March 14 th , 2011	Hydrogen explosion in unit 3 building.	
March 15 th , 2011	Fire reported in unit 4. Hydrogen explosion in unit 2.	
March 16 th , 2011	Fire reported in unit 4. Steam observed from unit 3.	
March 17–18 th , 2011	First use of helicopters to dump seawater on unit 3.	
March 20–21 st , 2011		French media report arrival of a radioactive cloud on March 23 rd
March 24 th , 2011		Radioactive Iodine 131 detected in center of France
March 25 th , 2011		Iodine 131 and Cesium 137 above detection thresholds
...		
April 25 th , 2011		Radioactive levels below detection thresholds.
...		
December 16 th , 2011	All reactors officially in state of cool shutdown.	

doi:10.1371/journal.pone.0058385.t002

the delivery of KI pills favors this hypothesis, at least for KI pills. The Fukushima crisis may have acted as a reminder to people authorized to stockpile to renew their stock. But it is clear that the control surrounding KI delivery in France prevented a wider unnecessary and potentially dangerous use. In the US where KI is available without prescriptions, several media sources reported that KI manufacturers and retailers were out of stock in the days following the disaster. However no study has described the extent of the phenomena and its public health impact yet.

Habits of self-medication are deeply rooted in the population of many countries. Combined with misinformation and distrust for official recommendations, it may lead, in time of public health crisis, to dramatic consequences. In the future, public health authorities may have to adapt their reaction accordingly to account for behaviors documented in this study. Consequently, it is crucial for public health decision makers to understand how and when populations resort to self-medication in order to design more efficient public health campaigns in times of public health crisis. Our study seems to show that real-time monitoring of drug sales could be an appropriate tool in that context.

References

1. Nuclear Emergency Response Headquarters Government of Japan (2011) Report of Japanese Government to the IAEA Ministerial Conference on Nuclear Safety - The Accident at TEPCO's Fukushima Nuclear Power Stations.
2. Nuclear and Industrial Safety Agency (2011) INES (the International Nuclear and Radiological Event Scale) Rating on the Events in Fukushima Dai-ichi Nuclear Power Station by the Tohoku District - off the Pacific Ocean Earthquake.
3. Yablokov AV, Nesterenko VB (2009) Chapter I. Chernobyl Contamination?: An Overview 1. Chernobyl Contamination through Time and Space. *Power* 30: 4–30.
4. Carlé B, Charron S, Milochevitch A, Hardeman F (2004) An inquiry of the opinions of the French and Belgian populations as regards risk. *Journal of hazardous materials* 111: 21–27.
5. Raude J, Caille-Brillet A-L, Setbon M (2010) The 2009 pandemic H1N1 influenza vaccination in France: who accepted to receive the vaccine and why? *PLoS currents* 2: RRN1188.
6. Setbon M, Le Pape M-C, Létroublon C, Caille-Brillet A-L, Raude J (2011) The public's preventive strategies in response to the pandemic influenza A/H1N1 in France: Distribution and determinants. *Preventive medicine* 52: 178–181.
7. Vaux S, Van Cauteren D, Guthmann J-P, Le Strat Y, Vaillant V, et al. (2011) Influenza vaccination coverage against seasonal and pandemic influenza and their determinants in France: a cross-sectional survey. *BMC public health* 11: 30.
8. Institut de Radioprotection et de Sûreté Nucléaire (n.d.). Available: <http://www.irsn.fr/EN/news/Pages/news.aspx>. Accessed 3 March 2012.
9. Autorité de Sûreté Nucléaire (ASN) (n.d.) La distribution d'iode stable autour des centrales nucléaires françaises. Available: <http://www.distribution-iode.com>. Accessed 8 January 2013.
10. Celtipharm (n.d.) Panel Xpr-SO (c). Available: <http://www.xpr-so.com/pages/Public/Accueil.aspx>. Accessed 1 November 2011.
11. Kish JL (1965) *Survey Sampling*. Wiley & Sons.
12. Watkins RE, Eagleson S, Veenendaal B, Wright G, Plant AJ (2008) Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia. *BMC medical informatics and decision making* 8: 37.
13. Astagneau P, Ancelle T, editors (2011) *Surveillance épidémiologique*. Paris: Lavoisier.
14. Spallek L, Krille L, Reiners C, Schneider R, Yamashita S, et al. (2011) Adverse Effects of Iodine Thyroid Blocking: a Systematic Review. *Radiation protection dosimetry*: 1–11.

Conclusion

Our study reports, describes, and quantifies some preventive lay behaviors during an international public health crisis. Those reactions have implications in terms of public health policy because they consist of uncontrolled use of drugs outside of their original purpose. Consequently we advocate the use of real time drug sales monitoring for instantaneous evaluation of population behavior during a global public health crisis.

Supporting Information

Table S1 Details of groups of medication sales. (DOCX)

Acknowledgments

The authors thank Celtipharm for graciously providing medication sales data.

Author Contributions

Conceived and designed the experiments: PC MP AB. Performed the experiments: PC. Analyzed the data: PC MP AB. Contributed reagents/materials/analysis tools: PC. Wrote the paper: PC AB.

2) Impact potentiel de la pandémie de grippe et de l'utilisation des gels hydro-alcooliques sur l'épidémie de gastro-entérite en France.

Ce travail a fait l'objet d'une publication scientifique en octobre 2013 dans la revue PLoSOne:

P. Crépey, M. Pivette, M. Desvarieux, "Potential Impact of Influenza A / H1N1 Pandemic and Hand-Gels on Acute Diarrhea Epidemic in France," *PLoS One*, vol. 8, no. 10, pp. 1–6, 2013.

a) Résumé

La pandémie H1N1 en 2009 a suscité de nombreuses attentions des autorités sanitaires qui ont mis en place des campagnes de vaccination et des mesures de prévention, recommandées par l'OMS. L'objectif de cette étude était d'analyser si cette pandémie et les mesures de santé publique mises en œuvre en France ont pu avoir un impact sur une autre maladie infectieuse, la gastro-entérite, dont l'évolution et l'historique sont connus.

Les données d'incidence de gastro-entérites attendues en 2009/2010 ont été estimées, en utilisant la méthode de bootstrapping appliquée aux données nationales de 2004 à 2009. Nous avons évalué les différences entre les incidences attendues et observées en 2009/2010. Nous avons également analysé les ventes de gels hydro-alcooliques pendant la même période.

Le nombre de cas de gastro-entérite en 2009/2010 étaient significativement inférieurs à ceux attendus jusqu'à la troisième semaine de décembre (-24%, 95% CI [-36%; -9%]), puis significativement supérieurs (+40%, 95% CI [22%; 62%]), entraînant globalement un surplus de 574440 cas par rapport aux années précédentes. L'épidémie a été retardée de 5 semaines avec un pic 1,3 plus élevé que celui attendu. Les ventes de gels hydro-alcooliques pour les mains étaient inversement corrélées avec les incidences des deux maladies, grippe et gastro-entérite. Chez les plus de 65 ans, aucun excès de cas de grippe ni aucun rebond dans le nombre de cas de gastro-entérite n'a été observé, malgré un délai similaire dans l'apparition de l'épidémie saisonnière de gastro-entérite.

Nos résultats suggèrent qu'au moins une maladie épidémique a suivi une évolution temporelle inattendue en 2009/2010. La gastro-entérite aiguë a été contrôlée au début de la pandémie grippale pour tous les groupes d'âge, mais a présenté plus tard un pic plus élevé que celui attendu chez les moins de 65 ans. Le délai pour l'apparition de la gastro-entérite saisonnière peut être en partie attribuable à l'utilisation des gels hydro-alcooliques ainsi qu'à d'autres

comportements de prévention d'infection. Le rebond et la recrudescence de cas de gastroentérites plus tard dans la saison pourraient, eux, s'expliquer par une « interaction » entre les deux épidémies : fragilisée par la grippe, la population a été plus facilement touchée par la gastro-entérite. Cette étude écologique a montré l'intérêt de suivre les ventes de médicaments pour pouvoir identifier les comportements préventifs dans la population et évaluer leur impact.

b) Article

Potential Impact of Influenza A/H1N1 Pandemic and Hand-Gels on Acute Diarrhea Epidemic in France

Pascal Crépey^{1,2*}, Mathilde Pivette^{1,3}, Moïse Desvarieux^{1,4,5}

1 École des hautes études en santé publique Rennes, Sorbonne Paris Cité, Paris, France, **2**UMR EPV Emergence des Pathologies Virales –190, Aix-Marseille University, Marseille, France, **3** MAP5, Université René Descartes-Paris5, Paris, France, **4** Department of Epidemiology, Mailman School of Public Health, Columbia University, New York City, New York, United States of America, **5** U 738, Institut National de la Santé et de la Recherche Médicale, Paris, France

Abstract

Background: The 2009 A/H1N1 influenza pandemic has received a great deal of attention from public health authorities. Our study examines whether this pandemic and the resulting public health measures could have impacted acute diarrhea, a prevalent, highly transmissible and historically monitored disease.

Methods: Using augmentation procedures of national data for the previous five years (2004–2009), we estimated the expected timing and incidence of acute diarrhea in France in 2009–2010 and evaluated differences with the observed. We also reviewed national hand gels for the same period.

Findings: Number of episodes of acute diarrhea in France in 2009–2010 was significantly lower than expected until the third week of December (–24%, 95% CI [–36%; –9%]), then significantly higher (+40%, 95% CI [22%; 62%]), leading to a surplus of 574,440 episodes. The epidemic was delayed by 5 weeks with a peak 1.3 times higher than expected. Hand-gels sales inversely correlated with incidence of both influenza-like illness and acute diarrheal disease. Among individuals >65 yo, no excess cases of influenza and no excess rebound in acute diarrhea were observed, despite similar delay in the onset of the seasonal diarrheal epidemic.

Interpretation: Our results suggest that at least one endemic disease had an unexpected behavior in 2009–2010. Acute diarrhea seems to have been controlled during the beginning of the pandemic in all age groups, but later peaked higher than expected in the younger population. The all-age delay in seasonal onset seems partly attributable to hand-gels use, while the differential magnitude of the seasonal epidemic between young and old, concurrent for both influenza and acute diarrhea, is compatible with disease interaction.

Citation: Crépey P, Pivette M, Desvarieux M (2013) Potential Impact of Influenza A/H1N1 Pandemic and Hand-Gels on Acute Diarrhea Epidemic in France. PLoS ONE 8(10): e75226. doi:10.1371/journal.pone.0075226

Editor: Benjamin J. Cowling, University of Hong Kong, Hong Kong

Received: September 20, 2012; **Accepted:** August 13, 2013; **Published:** October 4, 2013

Copyright: © 2013 Crépey et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was funded by Ecole des Hautes Etudes en Santé Publique. No external funding used for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: pascal.crepey@ehesp.fr

Introduction

The first cases related to a novel strain of Influenza A/H1N1 appeared in Mexico within the first months of 2009, resulting in the first influenza pandemic of the 21st century [1]. During the international crisis that followed, a number of containment measures were taken at the national [2,3] and international levels [4]. While some measures were specific to influenza (e.g. mass vaccinations campaigns, recommended by WHO and initiated in October 2009 in many countries including France [5]), a number of non-pharmaceutical measures able to impact communicable disease transmission generally were also variably recommended across communities: restrictions on mass transport services and nonessential travels [4], social distancing interventions (closing schools and workplaces and limiting public gatherings) [4,6], and avoiding contact with the sick. At the national level, the most universally recommended two-set measures were frequent hand-washing and availability and use of alcohol-based hand sanitizers; as well as covering sneezes/coughs, and the use of masks, all

widely advocated through extensive print, audio, video campaigns conducted throughout virtually all public spaces, airwaves and outlets. Recent theoretical models confirm the legitimacy of these early, sustained, and non-pharmaceutical interventions for influenza containment [7–12]. However, there is a severely lacking amount of information about the potential indirect impact of this international crisis on other prevalent diseases. The aim of this study was to take advantage of the unique natural experiment of large-scale multi-focal public health interventions and to assess their impact on a historically monitored commonly transmitted infectious disease.

Methods

We hypothesized that the collateral effect of the previously described large-scale interventions would reduce transmission of acute diarrheal disease in France.

Acute gastro-enteritis was chosen because it is among the most common illnesses worldwide, with more than 4.5 M cases in

France in 2011 (data from French *Sentinelles* Network). It is also highly transmissible with two main etiologic agents, rotavirus for children <2 years and Norovirus for adults [13], requiring small infectious doses [14].

Our analysis is based on syndromic data from the *Sentinelles* network database [15–17]. *Sentinelles* is a network of 1300 volunteer general practitioners working in metropolitan regions throughout France (see <http://www.sentiweb.org/>). They continuously report cases of eight different diseases. In particular, the database includes weekly reports of influenza like illness since 1984 and of acute diarrhea cases since 1990. A case of syndromic acute diarrhea was defined as a patient having at least three daily, soft or watery stools in the past 14 days. In order to analyze influenza incidence in France, we used the estimated number of Influenza-like illness cases (ILI, defined as patients with sudden fever >39°C with myalgia and respiratory signs) reported by the *same* network. This database has recently been used to show the impact of school closures –another kind of preventive measure– on influenza epidemic [17]. We acknowledge that ILI is a poor estimator of true influenza incidence. However, we are more interested in this study by the timing of the epidemic and by yearly comparisons than by the true level of influenza incidence.

We then estimated the expected weekly incidence, or number of cases, of acute diarrhea, and influenza during the 2009–2010 period. Data augmentation procedures (bootstrapping) were used to generate 1000 new datasets of weekly incidence, or number of cases, based on the five previous years (2004–2009). A 5% random noise was added to the data to account for potential inaccuracy in data collection. Thanks to this method we were able to compute a weekly average incidence and 95% bootstrap confidence intervals. Data interpolation was performed using a sliding average to remove noise in the data and highlight the general trend. This simple approach can be qualified as a non-adaptive version of the “*analogues method*” previously used on similar data [18], i.e. based on the basic premise that an epidemic is likely to behave as previous ones. To validate the method we estimated the 2008–2009 “normal” season of acute diarrhea epidemic using preceding years.

Observed and expected 2009/2010 values of incidence rates, number of cases, time taken to reach epidemic peak, and peak height were compared. The epidemic period was determined using the Serfling method [19]. The Serfling method uses cyclic regression to model seasonal pattern of incidence and to define epidemic threshold. We also conducted separate analysis for the >65years-old.

We chose to analyze alcohol-based hand gels because of their sustained recommendation at the national level, and their potential to interfere with fecal-oral transmission via chemical cleansing of the hands. Analysis of the evolution of hand gels sales in France is based on data from a stratified sample of 3004 pharmacies on the French metropolitan territory set up by the company Celtipharm [20]. These pharmacies automatically report their sales continuously several times a day since 2007. Thanks to a regularly updated exhaustive database of the 22,458 active French pharmacies, a stratification to improve representativeness is performed on sales revenue (6 levels for global revenue and 4 levels per type of sales: prescribed drugs, OTC, and other type of sales), localization (5 geographic areas) and sales area (5 types, from rural to densely urban). Each stratum has a minimum of 30 pharmacies or is merged with neighboring strata. Sampling rates per strata are computed with the Neyman optimal allocation algorithm [21]. Extrapolations from the sample have been validated with data from drug manufacturers who distribute their products directly (and only) to pharmacies.

Results

The 2009 influenza pandemic in France was different from previous seasonal influenza epidemics, both in timing and magnitude (Figure 1). The epidemic started in July 2009, with a peak in the middle of November, and ended in January 2010. In previous years (2004 to 2009), the epidemic lasted from November to April, with a peak in January (Figure S1). The 2009 epidemic peak was also higher than during previous seasonal influenza epidemics.

In 2009–2010, the number of acute diarrhea episodes in France was significantly lower than expected until the third week of December (–24%; 95% CI: –36%, –9%), and then significantly higher (+40%; 95% CI: 22%, 62%), leading to an excess of 574,440 episodes (+14%; 95% CI: –2%, +34%) (Table 1). As shown in Figure 2, the epidemic started the second week of December and was delayed by five weeks; the maximum incidence occurred during the first week of January, one week later than expected, with the peak 1.3 times higher. Details of the differences between estimated and observed incidence are in Table 1. Figure 3 and S2 show that the beginning of the acute diarrhea epidemic in previous years (2004–2009) was observed 3 weeks *before* the beginning of the influenza epidemic, with an epidemic peak one month before the peak of influenza. In 2009–2010, however, the acute diarrhea epidemic started right *at the end* of the influenza epidemic, with a peak six weeks after the peak of influenza.

Estimations of the 2008–2009 “normal season” using the same method based on 2004–2008 years or using a larger dataset (2000–2008) were consistent with observations (start and peak of the epidemic in the 95% confidence interval, see figures S3 and S4 for details).

Because the 2009–2010 Influenza A/H1N1 primarily affected the young, sparing the old, we decided to separate the >65 years-old to determine whether both the delay in onset and the higher magnitude rebound were observed in this age group. While the onset of the acute diarrhea seasonal epidemic was delayed, no excessive rebound was observed among older subjects (Figure 2b): the number of episodes in 2009/2010 was significantly lower than expected until the last week of December (–19%; 95% CI: –25%, –8%), a deficit which was never recouped, leading to an overall reduction of 11% in the total number of cases.

Sales of hand gels peaked just before and during the very early beginning of the influenza epidemic but plummeted right before the real beginning of the epidemic peak (Figure 4). In addition, the

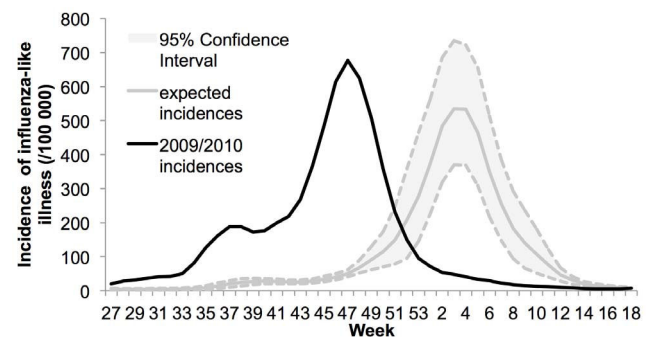


Figure 1. Incidences of influenza-like illness in 2009–2010 for France. Observed incidences are shown with black lines while expected incidences are shown with gray lines along their 95% confidence interval in dashed lines. Expected incidence is based on 2004–2009 data. The comparison shows the pandemic arrived sooner than previous years. The plot show smoothed data. doi:10.1371/journal.pone.0075226.g001

Table 1. Comparison between 2009–2010 observed acute diarrhea incidence and the expected incidence estimated from previous years (2004–2009) in France.

Differences with 2009–2010 incidence		low incidence	expected incidence	high incidence
Period between weeks 44 and 52	Incidence (/100 000)	–209	–608	–1,104
	Number of cases	–134,124	–391,297	–710,282
	% of increase or decrease	–9%	–24%	–36%
Period between weeks 53 and 18	Incidence (/100 000)	+2,019	+1,501	+935
	Number of cases	+1,298,442	+965,737	+601,107
	% of increase or decrease	+62%	+40%	+22%
Entire studied period between weeks 44 and 18	Incidence (/100 000)	+1,810	+893	–170
	Number of cases	+1,164,318	+574,440	–109,175
	% of increase or decrease	+34%	+14.1%	–2%

Low and high incidence scenarios are given by the bound of bootstrapped 95% confidence interval. The studied period is divided in two sub-periods, before and after the week 52 when the observed and expected incidences intersect.

doi:10.1371/journal.pone.0075226.t001

acute diarrhea epidemic started when the sales were at their lowest level. Of note, we also observe no hand-gels sales increase during the two epidemic seasons (2007–2008, 2008–2009) preceding the pandemic and the one following it (2010–2011) (Figure 4).

Discussion

We report unusual incidence patterns of acute diarrhea associated with the 2009–2010 influenza A/H1N1 pandemic and accompanying public health measures in France. Although we postulated a positive impact of non-pharmaceutical measures (notably alcohol-based hand-gels) on this communicable disease, the dramatic rebound in number of cases after the pandemic crisis was surprising. Our estimates translate into a 24% decrease in the number of acute diarrhea during the influenza pandemic followed

by a 40% increase afterwards, which corresponded to more than 570,000 additional diarrheal cases compared to previous years.

Thus, overall, we observed two phases: a disease control period during the influenza pandemic peak (translating into a delayed onset), then an epidemic rebound afterwards (translating into a higher peak magnitude). Regarding the first phase, at least two hypotheses can be offered. First, non-pharmaceutical interventions such as hand-gels may have had an impact on other communicable diseases transmitted by inter-individual contacts, thereby controlling acute diarrhea. Second, and perhaps concomitantly, competition between infections could have occurred, in which people already affected by influenza were more likely to reduce their contacts, hence affecting the pool of susceptibles for other communicable diseases. Consequently, the rebounding second stage may have coincided with a letdown in the preventive

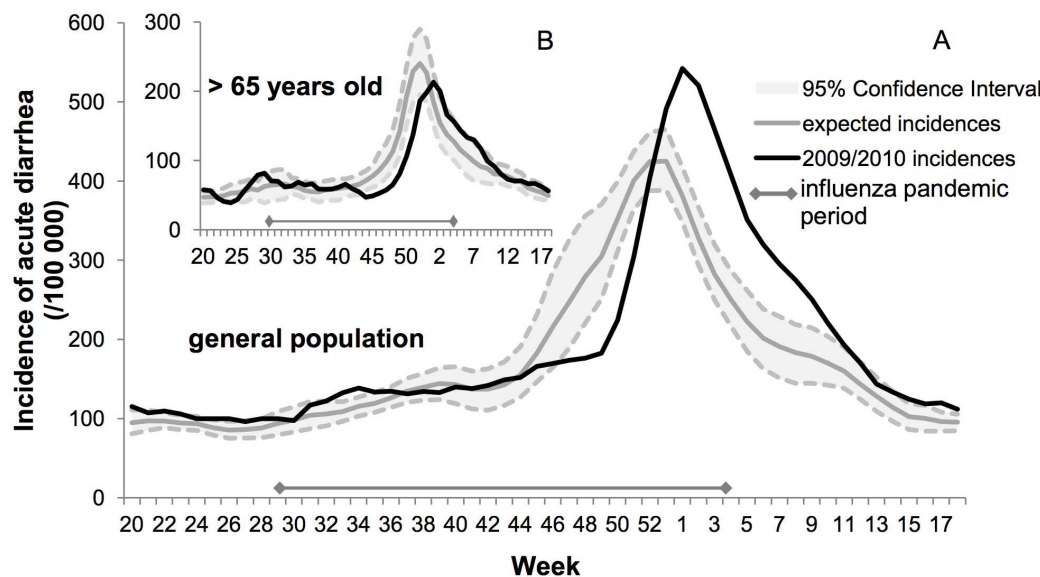


Figure 2. Incidences of acute diarrhea in France in 2009–2010. a) Observed incidence of acute diarrhea in general population (black line) peaks later and higher than the corresponding expected incidence (gray line, 95% confidence interval shown in dashed lines). b) In patients >65 years old, the peak is also delayed but its maximum is lower than expected. The interval shown at the bottom of the plots stands for the pandemic period in France defined by the Serfling method. Both plots show smoothed data.

doi:10.1371/journal.pone.0075226.g002

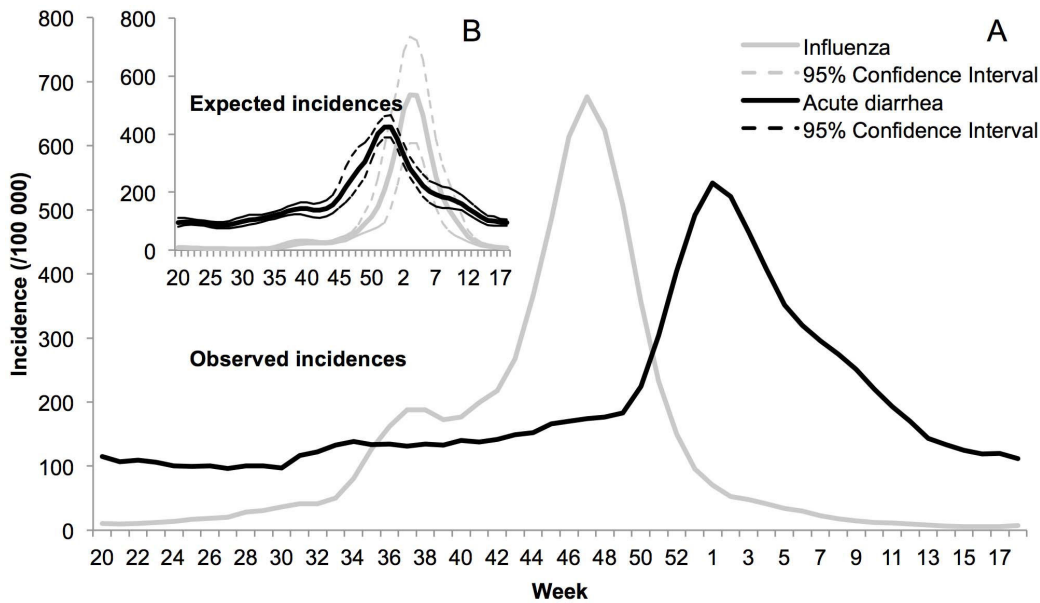


Figure 3. Comparison between influenza (gray line) and acute diarrhea (black line) incidence patterns in France. a) Observed incidences in 2009–2010 show that influenza epidemic preceded the acute diarrhea epidemic. During the preceding years the acute diarrhea epidemic was arriving first, as shown on **b)** with the expected incidences based on 2004–2009 data. Both plots show smoothed data.
doi:10.1371/journal.pone.0075226.g003

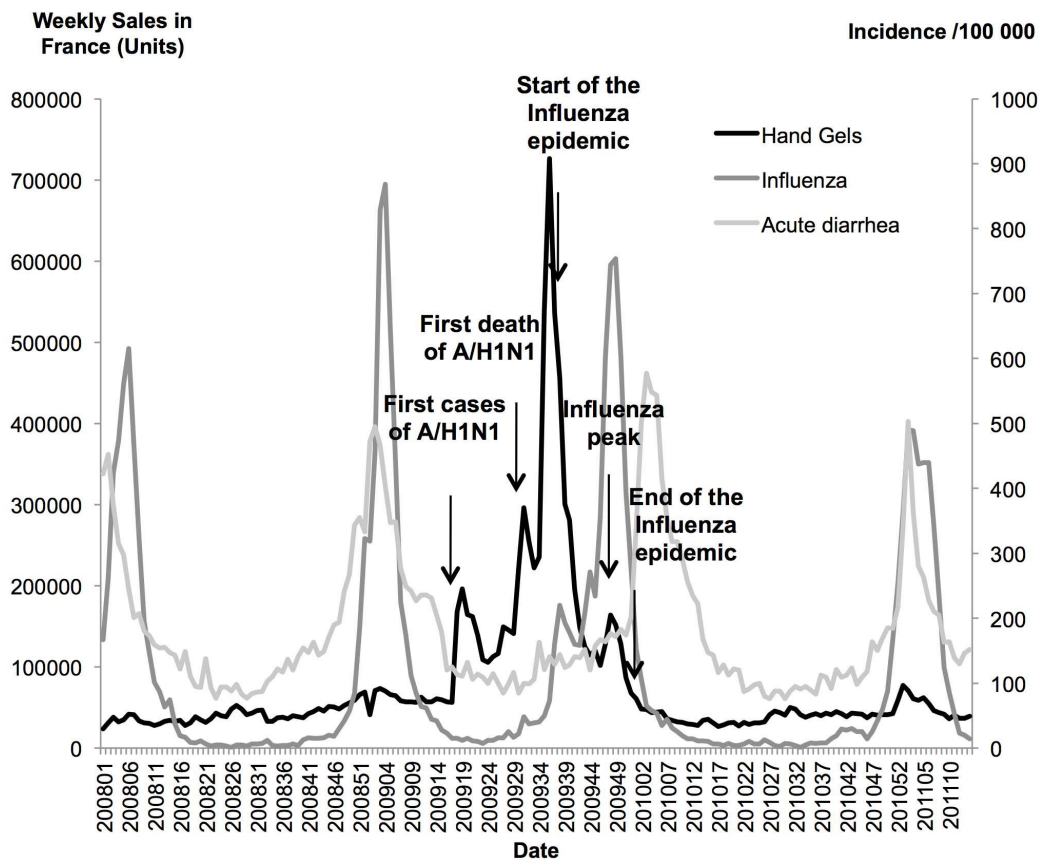


Figure 4. Weekly incidences of acute diarrhea and influenza-like-illness, along with weekly sales of hand gels in France from January 2008 to March 2011. We observe inverse correlations between sales of hand-gels, and incidence of influenza-like-illness and acute diarrhea. We also observe no hand-gels sales increase during the two epidemic seasons (2007–2008, 2008–2009) preceding the pandemic and the one following it (2010–2011).
doi:10.1371/journal.pone.0075226.g004

measures that were used in the first stage. An increased seasonal effect could also explain the higher magnitude of the later acute diarrhea epidemic: the epidemic arriving later in the winter season may have benefited from a higher impact of the weather on the force of infection (lower temperature, higher precipitation, etc...).

Alternatively, however, the population may have also been weakened by the influenza pandemic since this particular strain has been shown to trigger impaired immune response which may result in higher susceptibility [22]. However data from our surveillance network show that the influenza incidence rate in subjects 65 years and older was no different than in previous years [23]. In addition it seems that aged individuals developed a milder form of the disease compared to previous years and contrary to the younger population [24]. This population is critical in differentiating the impact of universally recommended public health measures from inter-disease interaction. Indeed, subjects older than 65 years old in France only experienced the delayed onset (potentially consistent with increased hand-gels use), but not the higher-than-expected rebounding stage in acute diarrheal disease. This discrepancy between delay with no higher magnitude, specific to that age group less affected than the younger population, could reveal a link between these two diseases in the young population, beyond a let down of public health measures. Not being as affected as the younger, the older population had no acute diarrhea rebound. The interrelated nature of diseases has been proposed several times [25–27] to explain for example the relationship between the 1918 influenza pandemic and declining death rates of tuberculosis, cardiovascular disease and nephritis in the US [28,29]. Here, excess cases of acute diarrhea could be explained as a “temporary harvesting effect” triggered by the influenza pandemic. All the above could fit diseases interaction models of syndemics [30] or pathocenosis [31], defined as the idea of complete interaction between diseases in an interrelated ecosystem.

Reporting bias could also have affected our data. An underreporting of acute diarrhea during the influenza pandemic could explain the sharp increase in the number of cases after the crisis. Furthermore, differential diagnosis could have been a major issue considering 38% of patients with influenza A/H1N1 presented with symptoms of nausea and diarrhea [32]. An alternative explanation could be the priorities, explicit or implicit, in which diseases were given treatment during the pandemic. In contrast, an over-reporting of acute diarrhea after the pandemic could partly explain the higher observed peak. For example, an increase in care seeking behavior may have occurred after the pandemic. However, these biases could explain some amplitude variations but not the observed epidemic delay of 6 weeks nor is there any easily identifiable reason why they would have spared the >65 years-old, all the more vulnerable to diarrhea-induced dehydration.

Hand gel sales analysis may present different bias. First, we don't know whether people who bought hand gels effectively used it. Health seeking behaviors vary by demographic, social and economic factors. Population consulting a general practitioner and buying health products may differ. Moreover, it is difficult to link a number of hand gel sales to a number of patients. However, the sales analysis is useful to determine temporal evolution and to detect abnormal phenomena.

References

1. World Health Organization (n.d.) Pandemic (H1N1) 2009 - Updates. Available: http://www.who.int/csr/don/2009_04_24/en/index.html.
2. Secrétariat Général de la Défense Nationale (2009) Plan national de prévention et de lutte “Pandémie grippale.”
3. US Department of Health and Human Services (2005) HHS Pandemic Influenza Plan.

Conclusion

Our results tend to show that the 2009–2010 acute diarrhea epidemic has been delayed compared with previous years. This delay is temporally associated with a peak of hand-gel sales, which could be a proxy for other control measures, and with the influenza pandemic peak. However, the nature of their associations is still unclear since it could be causal, or simply fortuitous. But the temporal proximity, the social or possibly biological arguments emphasize the need for further studies to clarify the nature of the link. Its existence would highlight the complexity of predicting the behavior of an epidemic without consideration to the ecological niche since delaying an epidemic may not imply decreasing the number of cases.

Supporting Information

Figure S1 Incidences of influenza-like illness in 2009–2010 and in the previous year (2004–2009) for France. Observed incidences of influenza in 2009–2010 (black lines) arrived sooner than previous years. The plot shows smoothed data. (PDF)

Figure S2 Incidences of acute diarrhea in 2009–2010 and in the previous year (2004–2009) in France. Observed incidence of acute diarrhea in 2009–2010 (black line) peaks later and higher than the incidence from previous years. The plot shows smoothed data. (PDF)

Figure S3 Incidences of acute diarrhea in 2008–2009 for France. Observed incidences are shown with black lines while expected incidences are shown with gray lines along their 95% confidence interval in dashed lines. Expected incidence are based on 2000–2008 data. The comparison shows that estimations of the 2008–2009 incidences are consistent with observations. The plot shows smoothed data. (PDF)

Figure S4 Incidences of acute diarrhea in 2008–2009 for France. Observed incidences are shown with black lines while expected incidences are shown with gray lines along their 95% confidence interval in dashed lines. Expected incidences are based on 2004–2008 data. The comparison shows that estimations of the 2008–2009 incidences are consistent with observations. The plot shows smoothed data. (PDF)

Acknowledgments

The authors thank Celtipharm for graciously providing access to their pharmaceutical sales data (Xpr-SO©).

Author Contributions

Conceived and designed the experiments: PC MD. Performed the experiments: PC MP. Analyzed the data: PC MP MD. Contributed reagents/materials/analysis tools: PC MP. Wrote the paper: PC MP MD.

4. World Health Organization (2009) Pandemic influenza preparedness and response. Available: http://www.who.int/influenza/resources/documents/pandemic_guidance_04_2009/en/index.html.
5. World Health Organization (n.d.) WHO recommendations on pandemic (H1N1) 2009 vaccines. Available: <http://www.who.int/csr/disease/>.
6. Wu JT, Cowling BJ, Lau EHY, Ip DKM, Ho L-M, et al. (2010) School Closure and Mitigation of Pandemic (H1N1) 2009, Hong Kong. *Emerging Infectious Diseases* 16: 10–13.
7. Glass RJ, Glass LM, Beyeler WE, Min HJ (2006) Targeted social distancing design for pandemic influenza. *Emerging infectious diseases* 12: 1671–1681.
8. Germann TC, Kadau K, Longini IM, Macken CA (2006) Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences of the United States of America* 103: 5935–5940.
9. Longini IM, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, et al. (2005) Containing Pandemic Influenza at the Source: Supporting Online Material. *Population* (English Edition).
10. Ferguson NM, Cummings DA, Fraser C, Cajka JC, Cooley PC, et al. (2006) Strategies for mitigating an influenza pandemic. *Nature* 442: 448–452.
11. Ferguson NM, Cummings DA, Cauchemez S, Fraser C, Riley S, et al. (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 437: 209–214.
12. Kelso JK, Milne GJ, Kelly H (2009) Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza. *BMC public health* 9: 117.
13. Anderson EJ, Weber SG (2004) Review Rotavirus infection in adults. *The Lancet* 4: 91–99.
14. Glass RI, Parashar UD, Bresee JS, Turcios R, Fischer TK, et al. (2006) Rotavirus vaccines: current prospects and future challenges. *Lancet* 368: 323–332.
15. Valleron A-J, Bouvet E, Garnerin P, Ménarès J, Heard I, et al. (1986) A computer network for the surveillance of communicable diseases: the French experiment. *American journal of public health* 76: 1289–1292.
16. Flahault A, Blanchon T, Dorléans Y, Toubiana L, Vibert JF, et al. (2006) Virtual surveillance of communicable diseases: a 20-year experience in France. *Statistical methods in medical research* 15: 413–421.
17. Cauchemez S, Valleron A-J, Boëlle P-Y, Flahault A, Ferguson NM (2008) Estimating the impact of school closure on influenza transmission from Sentinel data. *Nature* 452: 750–754.
18. Viboud C (2003) Prediction of the Spread of Influenza Epidemics by the Method of Analogues. *American Journal of Epidemiology* 158: 996–1006.
19. Serfling RE (1963) Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public health reports* 78: 494–506.
20. Celtipharm (n.d.) Panel Xpr-SO (c). Available: <http://www.xpr-so.com/pages/Public/Accueil.aspx>. Accessed 1 November 2011.
21. Kish JL (1965) *Survey Sampling*. Wiley & Sons.
22. Fox A, Hoa LNM, Horby P, Van Doorn HR, Trung NV, et al. (2012) Severe Pandemic H1N1 2009 Infection Is Associated with Transient NK and T Deficiency and Aberrant CD8 Responses. *PLoS one* 7: e31535.
23. Weil-Olivier C (2010) [First season of pandemic influenza A/H1N1]. *La Revue du praticien* 60: 1388–1391, 1394–1398.
24. Lemaitre M, Carrat F, Rey G, Miller M, Simonsen L, et al. (2012) Mortality burden of the 2009 A/H1N1 influenza pandemic in France: comparison to seasonal influenza and the A/H3N2 pandemic. *PLoS one* 7: e45051.
25. McNeill WH (1998) *Plagues and peoples*. Anchor Books.
26. Crosby AW (2003) *The Columbian exchange: biological and cultural consequences of 1492*. Greenwood Publishing Group.
27. Noymer A (2010) *Epidemics and Time, Influenza and Tuberculosis during and after the 1918–1919 Pandemic*. *Plagues and Epidemics: Infected Spaces Past and Present*. 137–152.
28. Noymer A, Garenne M (2000) The 1918 influenza epidemic's effects on sex differentials in mortality in the United States. *Population and Development Review*.
29. Garenne M, Noymer A (2008) Les effets à long terme de la grippe espagnole de 1918. *Cahiers de sociologie et de démographie médicales*.
30. Singer M, Clair S (2003) Syndemics and public health: reconceptualizing disease in bio-social context. *Medical anthropology quarterly* 17: 423–441.
31. Grmek MD (1969) Préliminaires d'une étude historique des maladies. *Annales Histoire, Sciences Sociales* 24: 1473–1483.
32. Team NS-OIA (H1N1) VI (2009) Emergence of a Novel Swine-Origin Influenza A (H1N1) Virus in Humans. *The New England journal of medicine* 360: 2605–2615.

3) Utilisation des médicaments psychotropes chez les enfants et adolescents en France

Ce travail a fait l'objet d'une publication scientifique en janvier 2015 dans la revue *Journal of Child and Adolescent Psychopharmacology*:

V. Kovess, S. Choppin, F. Gao, M. Pivette, M. Husky, and E. Leray. "Psychotropic Medication Use in French Children and Adolescents", *Journal of Child and Adolescent Psychopharmacology*, January 2015

a) Résumé

L'objectif de cette étude était de décrire la consommation de psychotropes chez les enfants et adolescents en France à partir des données de remboursement de l'Assurance Maladie en 2010. Les données proviennent d'un échantillon de 1% de la population des 0-17 ans (Echantillon Généraliste des Bénéficiaires). Les données recueillies contiennent l'âge, le sexe, les médicaments remboursés identifiés par leur code CIP et le type de prescripteur. Globalement, 2,5 % des enfants et adolescents ont eu une prescription contenant au moins un médicament psychotrope en 2010. La majorité des patients ont eu une prescription d'anxiolytiques (1,9%), suivies d'antidépresseurs (0,3%), d'antipsychotiques (0,3%) et de stimulants (0,2%). Entre 15 et 17 ans, 6,1% des filles ont eu une prescription d'anxiolytiques/hypnotiques et 1,1% une prescription d'antidépresseurs. Chez les garçons, les anxiolytiques/ hypnotiques sont les médicaments psychotropes les plus prescrits (1,8%). Ils consomment plus d'antipsychotiques que les filles entre 11 et 17 ans, et plus de psychostimulants que les filles entre 6 et 14 ans. La plupart des enfants ont reçu une prescription de la part d'un médecin généraliste (81,7%). La prévalence de la consommation de psychotropes en France est beaucoup plus faible que celle observée aux Etats-Unis. Par rapport aux pays européens, la consommation d'antidépresseurs et d'antipsychotiques est similaire, la consommation d'anxiolytiques est supérieure et celle des psychostimulants est inférieure en France.

b) Apport des données de ventes de médicaments en pharmacies

Les données provenant de l'EGB prennent en compte uniquement les médicaments qui ont été remboursés aux assurés sociaux. Nous avons analysé, en complément des données de remboursement, les données de ventes de médicaments issues de l'échantillon de 3004

pharmacies de *Celtipharm* afin de quantifier l'utilisation des anxiolytiques et des hypnotiques achetés en dehors d'une prescription médicale. Nous avons ciblé cette classe de médicaments car elle contient des médicaments accessibles sans prescription (notamment la phytothérapie) alors que les autres classes de médicaments ne contiennent que des médicaments à prescription obligatoire. Nous avons analysé la part de médicaments anxiolytiques et hypnotiques achetés sans ordonnance en 2010 chez des enfants de 0 à 17 ans. Les résultats montrent que 28% des médicaments anxiolytiques et hypnotiques ont été achetés sans ordonnance chez les 0-17 ans, cette proportion est de 35% chez les 0-2 ans. 80% de ces médicaments achetés sans ordonnance sont des produits de phytothérapie. Les résultats montrent que la consommation d'anxiolytiques et d'hypnotiques est sous-estimée en prenant en compte uniquement les données de remboursements et que les médicaments achetés sans ordonnance représentent une part importante de la consommation de psychotropes chez les enfants, notamment chez les plus jeunes. Il s'agit de données complémentaires aux données de remboursement dans ce type d'étude. Elles sont également utiles pour les comparaisons internationales, car les politiques de remboursement sont différentes d'un pays à l'autre et les comparaisons basées sur les remboursements peuvent sous-estimer la consommation réelle de ces médicaments.

VII. Discussion

1) Intérêt des données de ventes de médicaments par rapport aux données de surveillance existantes

Nos travaux ont montré que les données de ventes de médicaments présentent plusieurs avantages pour la surveillance des maladies infectieuses. En effet, l'obtention des données peut se faire en quasi temps-réel. Par exemple, les données issues de l'échantillon de pharmacies de *Celtipharm* sont accessibles avec un délai d'un jour, celles-ci sont ensuite consolidées dans les jours suivants. Cette rapidité d'obtention de données permet une évaluation quasi immédiate de l'état de santé de la population et, le cas échéant, autorise le déclenchement d'alertes en cas de détection d'événements inhabituels.

Ensuite, l'automatisation du recueil des données procure un avantage majeur au système par rapport aux outils de surveillance traditionnels basés sur le recueil actif d'information. Les réseaux de médecins sentinelles, les enquêtes en population, les déclarations obligatoires de

maladies par des cliniciens nécessitent du temps et l'effort des acteurs pour que la collecte des données puisse se faire.

Concernant le système que nous avons utilisé (échantillon *Celtipharm*), le volume de données recueilli est important car l'ensemble des ventes en officine est collecté et le nombre de pharmacies (13% des pharmacies) permet de couvrir une part importante de la population et de garantir une puissance importante.

De plus, les données collectées sont exhaustives, ce qui permet le suivi simultané de nombreuses pathologies. Les données de ventes sont non spécifiques mais elles sont cependant directement reliées à l'état de santé du patient, car c'est un acte d'achat de médicaments qui est enregistré, réalisé suite à la prescription d'un médecin ou au conseil d'un pharmacien. Ces données de surveillance sont donc plus spécifiques et informatives que des données sur les activités de recherches sur Internet (Ginsberg et al. 2009; Olson et al. 2013), sur les réseaux sociaux (Facebook, Twitter) ou l'absentéisme professionnel ou scolaire. Enfin, la collecte informatisée et automatique des données entraîne un coût plus faible par rapport à la mise en place d'enquêtes en population et la mise en place de réseaux de surveillance.

Les données de ventes peuvent apporter une information complémentaire aux données issues des bases médico-administratives, qui sont nécessaires pour identifier certains patients et décrire leurs caractéristiques. L'étude sur la consommation de médicaments psychotropes chez les enfants (Kovess et al. 2015) montre par exemple l'intérêt d'utiliser à la fois des données de remboursement et les ventes de médicaments achetés sans ordonnance pour estimer plus précisément l'incidence de maladies.

2) Détection d'épidémies

Les ventes de médicaments achetés sans prescription peuvent donner une information sur des pathologies légères ou sur les premiers symptômes d'une maladie, avant qu'elle ne nécessite une prise en charge médicale. Ainsi, nous avons montré que l'analyse des ventes de médicaments non-prescrits permettait de détecter les épidémies de gastro-entérites deux semaines plus tôt par rapport aux données issues des consultations médicales (Pivette et al. 2014b).

Cependant, il est nécessaire de s'attarder sur la spécificité et la sensibilité de ce type de signaux ainsi que sur le choix des médicaments « traceurs ». Ainsi, notre revue systématique de la littérature a montré que la plupart des études sur la détection d'épidémies étaient réalisées sur la grippe et la gastro-entérite, car ce sont des maladies qui disposent de données cliniques de surveillance et permettent de valider les signaux observés à partir des ventes de médicaments. Les données de ventes pourraient être utilisées pour la détection de maladies infectieuses non surveillées ou pour lesquelles les données de surveillance sont difficiles à obtenir ou coûteuses (maladies parasitaires, allergies, asthme...). Cependant, cette approche nécessite la collecte de données cliniques sur une période définie afin de valider les signaux observés.

La pertinence d'un système de surveillance basé sur les médicaments dépend du choix des groupes de médicaments utilisés comme proxy de la maladie. Ils doivent être suffisamment sensibles car un choix trop restreint de médicaments peut empêcher l'identification de certaines épidémies, et suffisamment spécifiques car des regroupements trop larges de médicaments sont susceptibles de déclencher de fausses alertes. Le choix dépend du type de maladies. En effet, pour des pathologies traitées par des médicaments spécifiques, il est simple de sélectionner des médicaments traceurs (ex : benzylpenicillin pour la syphilis). A l'inverse, pour les pathologies traitées par des médicaments ayant diverses indications, il faut utiliser des méthodes de sélection permettant des regroupements cohérents de médicaments. Dans notre étude sur la détection des épidémies de gastro-entérites en France (Pivette et al. 2014b), nous avons fait le choix de sélectionner des groupes de médicaments susceptibles d'être achetés ou prescrits en cas de gastro-entérites en se basant sur les indications thérapeutiques des produits. D'autres méthodes de sélection ont été développées dans des études (Das et al. 2005; Cami et al. 2009; Pelat et al. 2010) pour identifier sans a priori parmi l'ensemble des classes de médicaments les plus pertinentes pour le suivi d'une pathologie. Dans ces études, les groupes de médicaments présentant un profil temporel similaire à l'incidence de la pathologie d'intérêt ont été sélectionnés. Ces méthodes présentent l'avantage d'identifier tous les médicaments présentant un profil similaire aux données de référence et de limiter les oublis dans la sélection. Cependant, certaines maladies peuvent être épidémiques lors de périodes similaires et cela entraîne un risque de sélection inadaptée de médicaments traceurs. Ainsi, une sélection basée sur les indications thérapeutiques des médicaments permet de limiter les erreurs de classification. De plus, ces méthodes a priori se basent sur des données cliniques de référence pour sélectionner les médicaments traceurs. Cependant,

certains médicaments peuvent être vendus sans ordonnance et présenter des dynamiques temporelles différentes des données cliniques de référence et risquent d'être exclus de la sélection.

3) Evaluation des mesures de santé publique

Nos travaux sur l'estimation de la couverture vaccinale ont montré l'intérêt du suivi des ventes de vaccins pour évaluer en temps réel la situation sanitaire de la population et vérifier si les programmes de vaccination fonctionnent bien. Le suivi des ventes de vaccins contre la grippe chez les personnes à risque peut permettre d'adapter les messages de communication au cours de la saison épidémique. Leur intérêt est aussi particulièrement intéressant pour les vaccins non remboursés par l'Assurance Maladie et pour lesquels aucune estimation de la couverture vaccinale n'est disponible, comme la vaccination contre le rotavirus. Le suivi des ventes de médicaments, de par la rapidité d'obtention des données, est un outil efficace pour évaluer l'impact de recommandations. Une étude basée sur les ventes de produits contraceptifs en pharmacies (Lemaitre et al. 2015) a par exemple permis d'évaluer l'impact des recommandations de l'ANSM en décembre 2012 sur l'utilisation des pilules contraceptives de 3^{ème} et 4^{ème} génération qui augmentaient les risques de thrombose. L'analyse des ventes a montré un report de la contraception vers d'autres méthodes, notamment vers les pilules de 2^{ème} génération, et un niveau global stable de la contraception avant et après l'annonce médiatique.

4) Limites

La surveillance basée sur les ventes de médicaments présente des limites qui doivent être prises en compte. Les données de ventes sont souvent enregistrées initialement pour des raisons commerciales ou administratives et non pour de la veille sanitaire. Ces données contiennent donc peu d'informations sur le patient et sur ses antécédents médicaux. On ne peut disposer au maximum que du sexe, de l'âge et éventuellement d'indices sur son statut socio-économique via son IRIS de résidence. Ainsi, nous n'avons pas connaissance du statut médical du patient (par exemple Affections de Longue Durée).

Le suivi dans le temps de la consommation médicamenteuse du patient peut aussi être inaccessible pour des raisons de protection de l'anonymat. Des données avec une identification du patient, comme celles issues de l'EGB, permettent ce suivi dans le temps et

une description des trajectoires de soins. Les ventes de médicaments en pharmacies ne permettent pas non plus d'identifier des pathologies traitées selon d'autres modalités que la délivrance en officine (médicaments délivrés à l'hôpital, en centre de vaccination, PMI etc.). Certains vaccins sont par exemple administrés pour une grande part en PMI. C'est le cas pour 20% des ventes de vaccins BCG (Guthmann et al. 2012). L'estimation de la couverture vaccinale est donc plus robuste si elle est calculée via l'analyse des certificats de santé des enfants ou les ventes de vaccins aux collectivités. Il n'est pas non plus possible d'identifier des pathologies lourdes dont le traitement s'effectue essentiellement en milieu hospitalier, comme la tuberculose (Girard et al. 2014). Celles-ci peuvent néanmoins être surveillées via l'analyse des données du PMSI.

Les données de ventes ne sont pas non plus adaptées pour identifier des cas individuels d'une pathologie ou surveiller des maladies rares. En effet, même si les ventes issues d'un échantillon de pharmacies peuvent concerner une part importante de la population, certaines pathologies, si elles sont rares, n'auront qu'une faible probabilité d'être prises en compte. Celles-ci seront néanmoins enregistrées dans des bases de remboursements de l'assurance maladie qui, elles, sont exhaustives (DCIR).

Les ventes peuvent être influencées par des facteurs autres que l'état de santé de la population comme les promotions sur des produits, les périodes de vacances scolaires ou des annonces médiatiques. Les médicaments peuvent également être achetés pour traiter une maladie ou en prévention d'une maladie que le patient risque d'avoir (Miller et al. 2006). Il est ainsi nécessaire de différencier les produits achetés en traitement de maladies et ceux qui reflètent des comportements préventifs de santé de la population, comme cela a été le cas pour les achats de produits iodés pendant la crise de Fukushima en France (Crépey et al. 2013). L'achat et la consommation de médicaments varient également en fonction de facteurs tels que l'âge, le sexe, ou le niveau social (Frosst et al. 2006) et qui doivent être pris en compte dans l'interprétation des résultats.

Enfin, les ventes de médicaments sont collectées par des sociétés privées (Celtipharm, IMS,...) et leur accès, en dehors de collaboration permettant l'accès libre aux données, peut être coûteux et représenter un frein pour leur utilisation dans un cadre de recherche ou de mise en place de systèmes de surveillance.

5) Nouvelles sources de données

De nouvelles données de surveillance se développent grâce à l'accès à des volumes importants de données reliées à la santé, notamment grâce à l'essor du numérique. Google Flu trend est par exemple une application apparue en 2008 et présente dans 25 pays (Google). Son objectif est de suivre et détecter la grippe saisonnière à partir des requêtes réalisées sur Google par les internautes (Ginsberg et al. 2009; Carneiro and Mylonakis 2009). De nombreuses études ont montré de fortes corrélations entre les données de Google Flu Trend et des données de référence cliniques (Pelat et al. 2009; Patwardhan and Bilkovski 2012). L'analyse des données de réseaux sociaux se développent également (facebook, twitter). L'agence gouvernementale britannique (Food Standards Agency) a par exemple analysé les termes avec hastags (#) concernant la gastro-entérite sur Twitter pendant l'hiver 2012-2013 et a montré de fortes corrélations avec les notifications de norovirus venues de laboratoires. Néanmoins, ces données présentent des limites, les recherches sur Internet et les communications sur les réseaux sociaux peuvent être influencées par les médias (Cooper et al. 2005), les données ne permettent pas de définir de façon précise l'état de santé de la personne effectuant la recherche et aucune caractéristique sur le patient n'est disponible (Olson et al. 2013).

Une évolution observée dans les systèmes de surveillance est le développement de systèmes dans lesquels les patients déclarent eux-mêmes leur état de santé directement en ligne. Le site Influenzanet (Paolotti et al. 2014) permet par exemple de suivre l'activité grippale dans 10 pays d'Europe à partir de questionnaires remplis chaque semaine par des volontaires lors des saisons hivernales. Ces systèmes de surveillance sont réactifs et flexibles car l'information est directement transmise par Internet aux épidémiologistes sans intermédiaire. Cependant, il s'agit de données déclaratives et dont l'exactitude est difficile à démontrer. De plus, il existe des problèmes de représentativité de la population, les patients suivis étant des volontaires ayant fait la démarche de s'inscrire sur le site Internet. Une étude a également montré que certains patients étaient plus susceptibles de ne pas remplir régulièrement les questionnaires notamment les patients jeunes, les fumeurs ou les patients non-vaccinés contre la grippe (Bajardi et al. 2014). Ainsi, ces données sont certes réactives mais ne peuvent pas se substituer à une surveillance basée sur des données cliniques ou médicales.

En France, des projets de dossiers électroniques recensant des données médicales précises sur les patients ont été mis en place récemment et pourraient être des outils intéressants pour la surveillance épidémiologique.

Le dossier médical personnel est un projet lancé par le Ministère de la Santé en 2004 et dont la mise en place a commencé en 2011 (DMP). Il a pour objectif d'enregistrer de façon sécurisée les informations médicales du patient (compte-rendu d'hospitalisation, médicaments consommés, résultats d'examens, antécédents médicaux...). Il permet aux professionnels de santé, de partager, sous forme électronique les informations de santé concernant la prise en charge du patient. En janvier 2015, 500 000 personnes disposaient d'un Dossier Médical personnel en France. Le DMP permet d'avoir une vision complète de la situation médicale du patient et de pouvoir étudier sa consommation de soins dans le temps. Il présente donc un intérêt important pour la mise en place d'études épidémiologiques (Quantin et al. 2006). Cependant, l'objectif principal de cet outil est de pouvoir coordonner les soins et d'améliorer la prise en charge du patient. Son usage à un niveau populationnel pour des études épidémiologiques est discuté et mais n'est actuellement pas mis en place.

Un autre dossier électronique « le Dossier Pharmaceutique (DP) » a été créé en 2007 et est géré par le Conseil national des Pharmaciens (Adenot 2007). Il recense, pour chaque bénéficiaire de l'assurance maladie qui l'accepte, tous les médicaments délivrés au cours des quatre derniers mois, qu'ils soient prescrits ou conseillés par un pharmacien. En février 2015, 99,2% des officines sont reliés au DP et 36 758 906 dossiers pharmaceutiques ont été créés (Ordre National des pharmaciens 2015a). Le DP a par exemple été utilisé pour étudier le report des prescriptions de pilules de 3^{ème} et 4^{ème} génération vers des pilules de 2^{ème} génération et l'utilisation de la contraception d'urgence en 2012. Le DP est actuellement utilisé pour étudier l'impact d'annonces faites par l'ANSM et le ministère lors de crises de santé publique sur la consommation de médicaments, mais il n'a pas été encore utilisé pour des études épidémiologiques avec un suivi de patients dans le temps. Il faut noter qu'une partie de la population seulement pourra être enregistrée car l'accord du patient est obligatoire pour la création de ces dossiers personnels.

6) Protection des données personnelles

Les données de ventes de médicaments en pharmacies contiennent des informations sur les patients, à la fois sur son identité (âge, sexe, localisation) et sur son état de santé grâce au suivi du type de médicaments consommés. L'analyse de ces données doit donc respecter certaines règles pour assurer la protection des données personnelles et le respect des libertés individuelles. Ainsi l'anonymisation des données est rendue obligatoire afin d'éviter l'identification de patients en respectant les recommandations de la CNIL.

L'ouverture des données de santé est un débat actuel qui porte notamment en France sur l'utilisation et l'accès aux données du SNIIRAM. Son accès est actuellement limité aux régimes d'assurance maladie et des organismes publics définis par des arrêtés (organismes de recherches, agences publiques) et n'est pas ouvert aux organismes à but lucratif. La restriction d'accès est défendue pour respecter la protection individuelle des individus, la confidentialité des données médicales et éviter les abus dans l'utilisation des données sensibles de santé. Le risque principal étant la ré-identification des patients grâce au recoupage de plusieurs sources de données. Les partisans de l'opendata qui est un principe selon lequel les données (celles recueillies par les organismes publics) doivent être disponibles pour tous (citoyens et entreprises) et réutilisables, prônent la transparence et la possibilité d'améliorer la surveillance épidémiologique grâce à une utilisation plus importante des bases de données médico-administratives. Une exploitation plus importante du SNIIRAM, qui est une des bases médico-économiques les plus importantes au monde, pourrait en effet permettre d'améliorer la qualité des soins, d'identifier les mésusages ou cibler des mesures de prévention. Plusieurs pays notamment les Etats-Unis, la Suède et le Royaume-Uni se sont engagés dans des démarches d'opendata en santé, le NHS au Royaume-Uni met par exemple à disposition en ligne des données mensuelles de prescriptions des médecins généralistes (Rowlingson et al. 2013). L'enjeu de ce débat en France sur l'ouverture des données est de pouvoir respecter les deux droits, à savoir rendre accessible aux citoyens les données de santé qui les concernent et proposer un cadre juridique stricte qui permet une protection rigoureuse de la vie privée et des règles de confidentialité.

VIII. Conclusion

Les études développées dans cette thèse ont permis de montrer différentes applications du suivi des ventes de médicaments en pharmacies pour la surveillance épidémiologique. Nous avons pu démontrer au travers de ces études qu'il s'agissait d'un outil valide et utile pour améliorer la surveillance épidémiologique en France. Son intérêt principal réside dans la rapidité d'obtention des données qui offre ainsi la possibilité de déclencher rapidement une alerte en cas d'épidémies ou d'évènements inhabituels. Il s'agit également d'un outil réactif pour évaluer l'impact de mesures de santé publique ou vérifier si des recommandations sanitaires sont respectées dans la population. L'obtention de données de consommation de médicaments non remboursés est également un atout supplémentaire par rapport aux systèmes

de surveillance classique, car cela permet d'obtenir des informations sur des pathologies légères ou les premiers symptômes d'une maladie.

Cet outil est intéressant mais présente des limites, souvent retrouvées dans les systèmes de surveillance indirects, qui sont principalement la difficulté à associer les ventes à un diagnostic précis de la maladie, l'absence de suivi dans le temps de la consommation des patients et le manque d'informations sur les caractéristiques du patient.

Les données de ventes utilisées dans nos travaux étaient extrapolées au niveau national et régional. Le nombre de pharmacies en France est important (21 614 pharmacies en 2015) et les pharmacies sont réparties de façon homogène sur tout le territoire. Ainsi, les ventes de médicaments pourraient être utilisées pour suivre l'état de la population ou émettre des alertes à un niveau local dans de futures études. Ce type d'analyse nécessiterait néanmoins d'avoir un échantillon plus important de pharmacies pour avoir des estimations précises au niveau local.

Nos travaux ont montré que des systèmes de surveillance basés sur la collecte automatique de données de ventes de médicaments disponibles en quasi temps-réel existent et sont déjà opérationnels. La véritable évolution viendra du développement de systèmes aussi réactifs que ceux présentés dans cette thèse mais où il sera possible d'avoir un suivi dans le temps de la consommation des patients, tout en respectant l'anonymat des données personnelles des patients. De tels systèmes permettront d'identifier plus spécifiquement les groupes de patients ayant des pathologies cibles et permettront de décrire leurs trajectoires de soins, tout en gardant les avantages d'un suivi en temps-réel. Une autre évolution pour améliorer la surveillance épidémiologique passera par la possibilité de coupler les bases de données de ventes ou de consommation de médicaments à d'autres données médicales comme les données d'hospitalisation, les antécédents médicaux, les résultats d'examens, les causes de décès et des données socio-économiques. Ces appariements permettraient d'avoir une vision plus complète de l'état de santé de la population.

Les données qui ont été utilisées dans cette thèse sont collectées par une société privée et utilisées dans le cadre d'un partenariat public-privé. On peut regretter que ce type de données permettant une obtention en temps-réel de l'information ne soient pas collectées et utilisées en routine par les autorités sanitaires pour la surveillance épidémiologique. En outre, comme nous l'avons décrit précédemment, des bases de données médico-administratives nationales très riches existent en France et un accès plus large et facilité aux chercheurs à ces bases

permettrait une meilleure exploitation de ces bases à des fins de surveillance et de recherche en santé publique.

Bibliographie

- Adenot I (2007) The pharmaceutical record. *Ann Pharm Fr* 65:325–30.
- AFIPA, Celtipharm (2014) Baromètre AFIPA 2013 des produits de selfcare. www.afipa.org.
- ANSM (2014) Analyse des ventes de médicaments en France en 2013.
- Antona D, Lévy-Bruhl D, Baudon C, et al (2013) Measles elimination efforts and 2008-2011 outbreak, France. *Emerg Infect Dis* 19:357–64.
- Aslam AA, Tsou M-H, Spitzberg BH, et al (2014) The reliability of tweets as a supplementary method of seasonal influenza surveillance. *J Med Internet Res* 16:e250.
- Assurance Maladie (2015) Le SNIIRAM. <http://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/sniiram/structure-du-sniiram.php>. Accessed 10 Jan 2015
- Astagneau P, Ancelle T (2011) *Surveillance épidémiologique*, Lavoisiers.
- Auvigne V (2014) Les concepts cartographiques de openhealth.fr. http://staging.openhealth.fr/Portals/1/download/EKI_CELTI_Gastro2013_NoteMethodoCarto_V03.pdf. Accessed 19 Feb 2015
- Bajardi P, Vespignani A, Funk S, et al (2014) Determinants of Follow-Up Participation in the Internet-Based European Influenza Surveillance Platform Influenzanet. *J Med Internet Res* 16:e78.
- Berger M, Shiao R, Weintraub JM (2006) Review of syndromic surveillance: implications for waterborne disease detection. *J Epidemiol Community Health* 60:543–50.
- Blin P, Dureau-Pournin C, Foubert-Samier A, et al (2014) Parkinson's disease incidence and prevalence assessment in France using the national healthcare insurance database. *Eur J Neurol* 22:464–471.
- Buehler JW, Berkelman RL, Hartley DM, Peters CJ (2003) Syndromic surveillance and bioterrorism-related epidemics. *Emerg Infect Dis* 9:1197–204.
- Buehler JW, Hopkins RS, Overhage JM, et al (2004) Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks Recommendations from the CDC Working Group. *MMWR Morb Mortal Wkly Rep* 53 (RR05):1–11.
- Burkom HS (2003) Development, adaptation, and assessment of alerting algorithms for biosurveillance. *Johns Hopkins APL Tech Dig* 24:335–342.
- Caille-Brillet A-L, Raude J, Lapidus N, et al (2013) Trends in influenza vaccination behaviours - results from the CoPanFlu cohort, France, 2006 to 2011. *Euro Surveill* 18:20628.

- Caisse Nationale d'Assurance Maladie (2014) Campagne de vaccination contre la grippe saisonnière. http://www.sante.gouv.fr/IMG/pdf/Dossier_de_presse_campagne_grippe_16-10-14.pdf.
- Caisse Nationale d'Assurance Maladie (2013) Campagne de vaccination contre la grippe saisonnière. http://www.sante.gouv.fr/IMG/pdf/DP_grippe_2013_081013.pdf.
- Cami A, Wallstrom GL, Fowlkes AL, et al (2009) Mining aggregates of over-the-counter products for syndromic surveillance. *Pattern Recognit Lett* 30:255–266.
- Carneiro HA, Mylonakis E (2009) Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clin Infect Dis* 49:1557–64.
- Caserio-Schönemann C, Bousquet V, Fouillet A, Henry V (2014) Le système de surveillance syndromique SurSaUd@. *BEH* 3-4:38–44.
- Celtipharm Panel Xpr-SO.
- Chardon O, Blanpain N (2010) Projections de population 2007-2060 pour la France métropolitaine. www.insee.fr. Accessed 15 Jan 2015
- Chen H, Zeng D, Yan P (2010) *Infectious Disease informatics- Syndromic Surveillance for Public Health and Bio-Defense*, Springer.
- Chen J-H, Schmit K, Chang H, et al (2005) Use of Medicaid prescription data for syndromic surveillance--New York. *MMWR Morb Mortal Wkly Rep* 54 Suppl:31–4.
- Cisar P, Bosnjak S, Maravic Cisar S (2010) EWMA Algorithm in Network Practice. *Int J Comput Commun Control* 5:160.
- Cook S, Glass R, LeBaron C, Ho M (1990) Global seasonality of rotavirus infections. *Bull World Health Organ* 68:171–7.
- Cooper CP, Mallon KP, Leadbetter S, et al (2005) Cancer Internet search activity on a major search engine, United States 2001-2003. *J Med Internet Res* 7:e36.
- Costagliola D, Flahault A, Galinec D, et al (1991) A routine tool for detection and assessment of epidemics of influenza-like syndromes in France. *Am J Public Health* 81:97–9.
- Crépey P, Pivette M, Bar-Hen A (2013) Quantitative assessment of preventive behaviors in France during the Fukushima nuclear crisis. *PLoS One* 8:e58385.
- Das D, Metzger K, Heffernan R, et al (2005) Monitoring over-the-counter medication sales for early detection of disease outbreaks--New York City. *MMWR Morb Mortal Wkly Rep* 54 Suppl:41–6.
- Davies GR, Finch RG (2003) Sales of over-the-counter remedies as an early warning system for winter bed crises. *Clin Microbiol Infect* 9:858–63.
- DMP Dossier Médical Personnel. www.dmp.gouv.fr. Accessed 14 Aug 2014

- DREES (2011a) L'état de santé de la population en France. Objectif 43: Couverture vaccinale de la population. http://www.sante.gouv.fr/IMG/pdf/Etat_sante-population_2011.pdf.
- DREES (2011b) L'état de santé de la population en France. Objectif 39: Vaccination contre la grippe des groupes à risque. http://www.sante.gouv.fr/IMG/pdf/Etat_sante-population_2011.pdf.
- Edge VL, Pollari F, Lim G, et al (2004) Syndromic surveillance of gastrointestinal illness using pharmacy over-the-counter sales. A retrospective study of waterborne outbreaks in Saskatchewan and Ontario. *Can J public Heal* 95:446–50.
- EphMRA European Pharmaceutical Marketing Research Association. www.ephmra.org. Accessed 1 Mar 2015
- Esparza-Aguilar M, Gastañaduy P a, Sánchez-Uribe E, et al (2014) Diarrhoea-related hospitalizations in children before and after implementation of monovalent rotavirus vaccination in Mexico. *Bull World Health Organ* 92:117–25.
- Farrington CP, Beale AD (1993) Computer-aided detection of temporal clusters of organisms reported to the Communicable Disease Surveillance Centre. *Commun Dis Rep CDR Rev* 3:R78–82.
- Farrington CP, Beale AD (1996) A statistical algorithm for the early detection of outbreaks of infectious disease. *J R Stat soc* 159:547–563.
- Flahault A, Blanchon T, Dorléans Y, et al (2006) Virtual surveillance of communicable diseases: a 20-year experience in France. *Stat Methods Med Res* 15:413–21.
- Fouillet A, Medina S, Medeiros H, et al (2014) La surveillance syndromique en France en 2014. *BEH* 3-4:75–80.
- Fourquet F, Desenclos J-C, Maurage C, Baron S (2003) Le poids médico-économique des gastro-entérites aiguës de l'enfant : l'éclairage du Programme de Médicalisation des Systèmes d'Information (PMSI). *Arch Pédiatrie* 10:861–868.
- French GPs Sentinelles Network (Réseau Sentinelles). <http://websenti.u707.jussieu.fr/sentiweb/>. Accessed 15 Jan 2015
- Frost GO, Majowicz SE, Edge VL (2006) Factors associated with the use of over-the-counter medications in cases of acute gastroenteritis in Hamilton, Ontario. *Can J public Heal* 97:489–93.
- Gagneur A, Nowak E, Lemaitre T, et al (2011) Impact of rotavirus vaccination on hospitalizations for rotavirus diarrhea: the IVANHOE study. *Vaccine* 29:3753–9.
- Ginsberg J, Mohebbi MH, Patel RS, et al (2009) Detecting influenza epidemics using search engine query data. *Nature* 457:1012–4.

- Girard D, Antoine D, Che D (2014) Epidemiology of pulmonary tuberculosis in France. Can the hospital discharge database be a reliable source of information? *Med Mal Infect* 44:509–514.
- Gleizes O, Desselberger U, Tatochenko V, et al (2006) Nosocomial rotavirus infection in European countries: a review of the epidemiology, severity and economic burden of hospital-acquired rotavirus disease. *Pediatr Infect Dis J* 25:S12–21.
- Google (Google Flu Trend). <https://www.google.org/flutrends/>. Accessed 15 Jan 2015
- Guthmann J, Fonteneau L, Lévy-bruhl D (2012) Mesure de la couverture vaccinale en France. Sources de données et données actuelles.
- Guthmann JP, Antoine D, Fonteneau L, et al (2011) Assessing BCG vaccination coverage and incidence of paediatric tuberculosis following two major changes in BCG vaccination policy in France. *Euro Surveill* 16:1–6.
- HCSP (2013) Vaccination des nourrissons contre les infections à rotavirus- recommandations.
- HCSP (2010) Vaccination systématique contre les rotavirus des nourrissons âgés de moins de 6 mois.
- Herzig P, Hartmann C, Fischer D, et al (1998) Pertussis complications in Germany--3 years of hospital-based surveillance during the introduction of acellular vaccines. *Infection* 26:227–31.
- Hogan WR, Tsui F-C, Ivanov O, et al (2003) Detection of Pediatric Respiratory and Diarrheal Outbreaks from Sales of Over-the-counter Electrolyte Products. *J Am Med Informatics Assoc* 10:555–562.
- INSEE (2015) Evolution et structure de la population. www.insee.fr. Accessed 15 Jan 2015
- INSEE Institut National de la Statistiques et des Etudes Economiques. <http://www.insee.fr/>. Accessed 10 Mar 2015
- Invs (2013) Surveillance syndromique à la Réunion. *Bull Veill Sanit* 21:3.
- Invs (2015) Epidémiologie de la grippe. <http://www.invs.sante.fr/Dossiers-thematiques/Maladies-infectieuses/Maladies-a-prevention-vaccinale/Grippe>. Accessed 15 Jan 2015
- Jackson ML, Baer A, Painter I, Duchin J (2007) A simulation study comparing aberration detection algorithms for syndromic surveillance. *BMC Med Inform Decis Mak* 7:6.
- Josseran L, Caillère N, Brun-Ney D, et al (2009) Syndromic surveillance and heat wave morbidity: a pilot study based on emergency departments in France. *BMC Med Inform Decis Mak* 9:14.

- Josseran L, Fouillet A, Caillère N, et al (2010) Assessment of a syndromic surveillance system based on morbidity data: results from the Oscour network during a heat wave. *PLoS One* 5:e11984.
- Kara EO, Elliot AJ, Bagnall H, et al (2011) Absenteeism in schools during the 2009 influenza A(H1N1) pandemic: a useful tool for early detection of influenza activity in the community? *Epidemiol Infect* 1–9.
- Kirian ML, Weintraub JM (2010) Prediction of gastrointestinal disease with over-the-counter diarrheal remedy sales records in the San Francisco Bay Area. *BMC Med Inform Decis Mak* 10:39.
- Kish JL (1965) *Survey Sampling*. Wiley & Sons
- Kovess V, Choppin S, Gao F, et al (2015) Psychotropic Medication Use in French Children and Adolescents. *J Child Adolesc Psychopharmacol* 25:168–175.
- Kusnik-Joinville O, Weill A, Salanave B, et al (2008) Prevalence and treatment of diabetes in France: trends between 2000 and 2005. *Diabetes Metab* 34:266–72.
- Langmuir AD (1963) The surveillance of communicable diseases of national importance. *N Engl J Med* 268:182–92.
- Lemaitre M, Lastennet G, Syr D, et al (2015) Impact of the 2013 French Pill Crisis on Women’s Behaviour Regarding Contraception. *Drugs - Real World Outcomes* 2:13–22.
- Lewis MD, Pavlin JA, Mansfield JL, et al (2002) Disease outbreak detection system using syndromic data in the greater Washington DC area. *Am J Prev Med* 23:180–6.
- Lombardo J, Burkom H, Elbert E, et al (2003) A systems overview of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE II). *J Urban Health* 80:i32–42.
- Magruder SF, Lewis SH, Najmi A, Florio E (2004) Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance. *MMWR Morb Mortal Wkly Rep* 53 Suppl:117–22.
- Martin-Latry K (2010) Pharmacoepidemiological research using French reimbursement databases : yes we can ! y. *Pharmacoepidemiol Drug Saf* 256–265.
- Miller J, Schmit K, Duncan P, Waters C (2006) Increased antiviral medication sales before the 2005-06 influenza season--New York City. *MMWR Morb Mortal Wkly Rep* 55:277–9.
- Ministère des Affaires sociales et de la Santé (2014) Calendrier des vaccinations et recommandations vaccinales 2014.
- Ministère des Affaires sociales et de la Santé (2013) Arrêté du 31 juillet 2013 modifiant la liste des spécialités remboursables aux assurés sociaux.

- Mostashari F, Fine A, Das D, et al (2003) Use of ambulance dispatch data as an early warning system for communitywide influenzalike illness, New York City. *J urban Heal* 80:i43–9.
- Mouly D, Vincent N, Guthmann J, Lévy-bruhl D (2013) Place du suivi des ventes de vaccins RRO pour évaluer l'impact d'une action de sensibilisation à la vaccination contre la rougeole réalisée en 2011 en Auvergne, France. *BEH* 8-9:77–82.
- Olson DR, Konty KJ, Paladini M, et al (2013) Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput Biol* 9:e1003256.
- Ordre National des pharmaciens (2015a) Nombre d'officines en France. <http://www.ordre.pharmacien.fr/Le-pharmacien/Secteurs-d-activite/Pharmacie/Cartes-regionales-Officine/Nombre-d-officines>. Accessed 15 Feb 2015
- Ordre National des pharmaciens (2015b) Nombre d'habitants par pharmacies en France. <http://www.ordre.pharmacien.fr/Le-pharmacien/Secteurs-d-activite/Pharmacie/Cartes-regionales-Officine/Nombre-d-habitants-par-officine>. Accessed 15 Feb 2015
- Paolotti D, Carnahan A, Colizza V, et al (2014) Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. *Clin Microbiol Infect* 20:17–21.
- Parez N, Allaert F, Derrough T, Caulin E (2007) Place et caractéristiques cliniques des gastroentérites aiguës à rotavirus chez les enfants de moins de cinq ans suivis en médecine de ville en France. Étude ROTAScore. *Pathol Biol* 55:453–459.
- Parez N, Mory O, Pozzetto B, et al (2012) Impact of Rotavirus gastroenteritis requiring hospitalization or presenting to emergency room among children less than 5 years in France. *Pathol Biol (Paris)* 60:275–81.
- Patwardhan A, Bilkovski R (2012) Comparison: Flu Prescription Sales Data from a Retail Pharmacy in the US with Google Flu Trends and US ILINet (CDC) Data as Flu Activity Indicator. *PLoS One* 7:e43611.
- Pelat C (2010) Nouveaux outils et nouvelles données pour la surveillance des maladies infectieuses. Université Paris VI
- Pelat C, Boëlle P-Y, Turbelin C, et al (2010) A method for selecting and monitoring medication sales for surveillance of gastroenteritis. *Pharmacoepidemiol Drug Saf* 19:1009–18.
- Pelat C, Turbelin C, Bar-Hen A, et al (2009) More diseases tracked by using Google Trends. *Emerg Infect Dis* 15:1327–8.
- Pivette M, Mueller JE, Crépey P, Bar-Hen A (2014a) Drug sales data analysis for outbreak detection of infectious diseases: a systematic literature review. *BMC Infect Dis* 14:604.
- Pivette M, Mueller JE, Crépey P, Bar-Hen A (2014b) Surveillance of gastrointestinal disease in France using drug sales data. *Epidemics* 8:1–8.

- Proctor ME, Blair KA, Davis JP (1998) Surveillance data for waterborne illness detection: an assessment following a massive waterborne outbreak of *Cryptosporidium* infection. *Epidemiol Infect* 120:43–54.
- Quantin C, Guinot C, Tursz A, et al (2006) Le traitement épidémiologique du Dossier Médical Personnel au service des malades ? *Rev Epidemiol Sante Publique* 54:177–184.
- Roberts S. (1959) Control chart tests based on Geometric Moving Averages. *Technometrics* 1:239–250.
- Rodman JS, Frost F, Jakubowski W (1998) Using nurse hot line calls for disease surveillance. *Emerg Infect Dis* 4:329–32.
- Rolland E, Moore KM, Robinson VA, McGuinness D (2006) Using Ontario’s “Telehealth” health telephone helpline as an early-warning system: a study protocol. *BMC Health Serv Res* 6:10.
- Rowlingson B, Diggle P, Taylor B, Lawson E (2013) Mapping English GP prescribing data: a tool for monitoring health-service inequalities. *BMJ Open* 3:e001363.
- Serfling RE (1963) Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Rep* 78:494–506.
- Simoens C, Sabbe M, Damme P Van, et al (2009) Introduction of human papillomavirus (HPV) vaccination in Belgium, 2007-2008. *Euro Surveill* 8:2–5.
- Stroup DF, Williamson GD, Herndon JL, Karon JM (1989) Detection of aberrations in the occurrence of notifiable diseases surveillance data. *Stat Med* 8:323–9.
- Sugawara T, Ohkusa Y, Ibuka Y, et al (2012) Real-time Prescription Surveillance and its Application to Monitoring Seasonal Influenza Activity in Japan. *J Med Internet Res* 14:e14.
- Tuppin P, Choukroun S, Samson S, et al (2012a) Vaccination against seasonal influenza in France in 2010 and 2011: decrease of coverage rates and associated factors. *Presse Med* 41:e568–76.
- Tuppin P, Cuerq A, Weill A, et al (2012b) Alzheimer disease and other dementia in France: Identification, management and neuroleptic use (2007-2009). *Rev Neurol (Paris)* 168:152–60.
- Tuppin P, Samson S, Weill A, et al (2009) Influenza vaccination coverage in France in 2007-2008: contribution of vaccination refund data from the general health insurance scheme. *Med Mal Infect* 39:780–8.
- Tuppin P, Samson S, Weill A, et al (2011) Seasonal influenza vaccination coverage in France during two influenza seasons (2007 and 2008) and during a context of pandemic influenza A(H1N1) in 2009. *Vaccine* 29:4632–7.

- Turbelin C, Pelat C, Boelle P-Y, et al (2009) Early estimates of 2009 pandemic influenza A(H1N1) virus activity in general practice in France: incidence of influenza-like illness and age distribution of reported cases. *Eurosurveillance* 14:1–3.
- Van Cauteren D, De Valk H, Vaux S, et al (2012) Burden of acute gastroenteritis and healthcare-seeking behaviour in France: a population-based study. *Epidemiol Infect* 140:697–705.
- Vaux S, Noël D, Fonteneau L, et al (2010) Influenza vaccination coverage of healthcare workers and residents and their determinants in nursing homes for elderly people in France : a cross-sectional survey. *BMC Public Health* 10:159.
- Watkins RE, Eagleson S, Veenendaal B, et al (2008) Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia. *BMC Med Inform Decis Mak* 8:37.
- Wohl H (1977) The cusum plot: its utility in the analysis of clinical data. *N Engl J Med* 296:1044–5.

Liste des tableaux et Figures

Tableaux

Tableau 1: Sensibilité, proportion de fausses alertes et délai de détection des épidémies de gastro-entérites à partir des ventes de médicaments non prescrits à partir de la méthode de détection EWMA.

Tableau 2: Vaccins contre la grippe disponibles en pharmacie lors de chaque saison hivernale (Nom commercial et code CIP du produit)

Tableau 3: Couverture vaccinale estimée lors de chaque saison hivernale (1^{er} septembre-31 mars). Population générale. France Métropolitaine.

Tableau 4 : Dates de début et de fin de campagne vaccinale (2008-2013)

Tableau 5 : Couvertures vaccinales (%) par groupe d'âge (plus de 65 ans), obtenues par les ventes de vaccins et les données de remboursements de la CNAM. 2009-2014, France Métropolitaine.

Figures

Figure 1 : Sources de données pour la surveillance épidémiologique (Invs 2013)

Figure 2 : Organisation du marché français du médicament.

Figure 3: Décomposition 2013 du marché officinal selon le statut des spécialités (ANSM 2014) A) En quantité B) En valeur

Figure 4 : Schéma du fonctionnement de la méthode CUSUM. La différence entre les valeurs observées et les valeurs attendues est calculée à chaque temps. La somme de ces différences depuis le temps 1, c'est-à-dire le CUSUM, est représentée par les barres bleues (Pelat 2010).

Figure 5: Illustration du choix d'une période de référence de 7 jours et d'une période tampon de 2 jours. L'observation courante est au temps 0.

Figure 6: Nombre de boîtes de médicaments pour le traitement de la gastro-entérite vendues sans ordonnance par semaine (ligne bleue) et Score de Cusum calculé chaque semaine (ligne rouge). France Métropolitaine. 2007-2012.

Figure 7 : Nombre de boîtes de médicaments sans ordonnance vendues pour le traitement de la gastro-entérite, par semaine entre 2007 et 2012, France Métropolitaine (données lissées avec un coefficient $\lambda = 0,3$). Les seuils de détection épidémiques définis à partir des coefficients $k=1$, $k=2$ et $k=3$ sont représentés par des lignes pointillées.

Figure 8 : Illustration de la méthode de Stroup pour la semaine s15 de 2015 à partir des valeurs des 5 années précédentes.

Figure 9 : Taux d'incidence hebdomadaire des syndromes grippaux en France (gris) et son intervalle de confiance, semaines 23 à 37 de 2009. Le taux d'incidence attendu (bleu) est calculé comme la moyenne des valeurs historiques (s-1 ; s ; s+1) des 24 années précédentes; le 95^{ème} quantile des observations historiques (bleu pointillé) est le seuil de détection. Le seuil est dépassé lors de la semaine 31 (Turbelin et al. 2009).

Figure 10: Représentation géographique de la gastro-entérite en France Métropolitaine (10/02/2015) à partir de l'indicateur développé par l'entreprise *Celtipharm*.

Figure 11 : Nombre de vaccins antigrippaux vendus en pharmacie, nombre de cas de grippe (Réseau Sentinelles) et périodes de campagne de vaccination. Semaine 17 2008 à semaine 35 2014, France Métropolitaine.

Figure 12: Taux cumulés journaliers de ventes de vaccins grippe lors de chaque campagne de vaccination. A) Vaccins prescrits B) Vaccins non-prescrits. Population générale, période de campagne vaccinale, 2008-2013, France métropolitaine

Figure 13 : Nombre de vaccins vendus par semaine entre la semaine 36 et la semaine 5, pour les saisons de 2009/2010 à 2013/2014. Population des ≥ 65 ans, France Métropolitaine (Hors Corse).

Figure 14: Couverture vaccinale cumulée journalière lors de chaque saison (1^{er} septembre- 30 mars) entre 2009/2010 et 2013/2014. Population des ≥ 65 ans, France Métropolitaine (Hors Corse).

Figure 15: A) Couvertures vaccinales à rotavirus estimées par an entre 2008 et 2013. B) Nombre de vaccins vendus par semaine (Rotarix / Rotateq). France Métropolitaine (Hors Corse).

Figure 16: Couverture vaccinale à rotavirus par région (2008-2012). France Métropolitaine (hors Corse).

Figure 17: Proportion de chaque type de gastro-entérites (%) et incidence totales des hospitalisations pour Maladies infectieuses et intestinales (A00-A09) (ligne rouge) entre 2008 et 2012 chez les enfants de 0 à 2 ans. France Métropolitaine, décembre à mai pour chaque saison.

Figure 18: Taux annuel moyen (pour 100000) d'hospitalisations pour Entérites à Rotavirus et Infections intestinales virales non précisées chez les enfants de 0 à 2 ans, par région, France Métropolitaine (2008-2102).

Figure 19 : Corrélation entre le taux annuel moyen d'hospitalisation (/100000) pour Entérites à Rotavirus et Infections intestinales virales non précisées chez les enfants de 0-2 ans et la couverture vaccinale moyenne à Rotavirus, par région, France Métropolitaine, 2008-2012. Chaque point représente une région.