



HAL
open science

Épidémiologie de la grippe en population générale à partir d'une cohorte participative

Caroline Guerrisi

► **To cite this version:**

Caroline Guerrisi. Épidémiologie de la grippe en population générale à partir d'une cohorte participative. Médecine humaine et pathologie. Sorbonne Université, 2019. Français. NNT : 2019SORUS128 . tel-02950849

HAL Id: tel-02950849

<https://theses.hal.science/tel-02950849>

Submitted on 28 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**SORBONNE
UNIVERSITÉ**

CRÉATEURS DE FUTURS
DEPUIS 1257

**THESE DE DOCTORAT DE
SORBONNE UNIVERSITE**

Spécialité Epidémiologie

Ecole Doctorale Pierre Louis de Santé Publique à Paris :
Epidémiologie et Sciences de l'Information Biomédicale

Présentée par

Mme Caroline GUERRISI

Pour obtenir le grade de

DOCTEUR de SORBONNE UNIVERSITE

**Epidémiologie de la grippe en population générale
à partir d'une cohorte participative**

soutenue le 20 novembre 2019

devant le jury composé de :

Mme Vittoria COLIZZA

Directrice de thèse

M. John PAGET

Rapporteur

M. Philippe TUPPIN

Rapporteur

M. Olivier STEICHEN

Examineur

Laboratoire de rattachement

**UMR-S 1136, Inserm | Sorbonne Université, Paris
Institut Pierre Louis d'Epidémiologie et de Santé Publique**

Directeur : Fabrice CARRAT

Equipe 1 : « Maladies transmissibles : Surveillance et Modélisation »

Responsable : Pierre-Yves BOËLLE

Réseau Sentinelles

Faculté de Médecine Sorbonne Université, site Saint-Antoine

27 rue Chaligny

75571 Paris cedex 12

France

Remerciements

Je tiens à remercier tout d'abord, Vittoria Colizza, qui a accepté de diriger cette thèse. Tu as partagé ton expérience qui m'a permis d'avancer dans ce projet et de m'ouvrir au monde de la recherche. Merci pour ton accompagnement dans ces travaux, ta confiance et ta disponibilité durant ces 3 ans.

Je remercie chaleureusement Jean-François Cosson, John Paget, Olivier Steichen et Philippe Tuppin de m'avoir fait l'honneur d'accepter de participer à mon jury de thèse.

Merci à Thomas Hanslik pour sa disponibilité et son soutien dans la réalisation de ce projet. Un merci particulier à Thierry Blanchon qui a su me convaincre d'entreprendre cette thèse. Je te remercie sincèrement de m'avoir aidé à faire ce choix. Ton expertise, ta disponibilité, ton soutien et ton optimisme m'ont beaucoup apporté durant ces 3 ans. Merci à Clément Turbelin pour les échanges réalisés autour de GrippeNet.fr. Tes avis m'ont toujours été d'une aide précieuse pour orienter les travaux à mener. Merci beaucoup à Ana et Cécile des bons moments partagés, de nos échanges studieux et moins studieux dans notre bureau. Je vous remercie sincèrement pour tout le soutien que vous m'avez témoigné au fil de ce projet. Un grand merci à Louise pour ses encouragements et sa confiance au quotidien. Enfin, un immense merci à toute l'équipe du réseau Sentinelles (présente et passée) pour leur bienveillance et bonne humeur, merci à : Marion, Mathieu, Jennifer, Caroline, Yves, Mathilde, Laetitia, Charly, Titouan, Corentin, Shirley, Natacha, Lore, Romain, Camille, Daouda, etc. C'est un réel plaisir de travailler à vos côtés.

Je tiens à remercier également l'équipe surveillance et modélisation de l'Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP), en particulier Pierre-Yves Boëlle pour sa disponibilité et les « italiens du 8^{ème} ». Plus largement, merci aux personnes des autres équipes et de l'administration de l'IPLESP avec qui j'ai pu partager quelques discussions, en particulier les personnes des 8^{ème} et 3^{ème} étages de St Antoine.

Un grand merci aux personnes impliquées de près ou de loin dans l'étude GrippeNet.fr, en particulier aux personnes de Santé publique France, aux co-auteurs des travaux menés, aux collègues des autres plateformes européennes, aux anciens membres de l'équipe, aux stagiaires. Merci pour vos idées et votre expertise dans la mise en place de la cohorte GrippeNet.fr et des projets attendants.

J'adresse un énorme merci à tous mes ami(e)s qui m'ont soutenu durant ces trois ans. Le manque de moments partagés à vos côtés durant la dernière ligne droite m'a motivée à accélérer la cadence de travail.

Je remercie chaleureusement toute ma famille et belle-famille. Merci sincèrement pour vos encouragements continus et vos attentions.

Merci du fond du cœur à mes parents. Merci pour votre soutien sans faille, pour tout ce que vous m'avez appris et permis d'apprendre. Vous m'avez toujours conforté et rassuré dans mes choix.

Merci infiniment à mes Sorelline pour leur écoute et tous les moments que l'on partage. Vous êtes mes piliers.

Enfin, le plus grand des mercis à Alban pour sa présence, son écoute et ses encouragements quotidiens. Tu m'as soutenue (et supportée) tout au long de ces trois années. Je te promets plein d'autres beaux projets à partager ensemble.

Résumé

L'émergence d'Internet et de la rapidité des échanges d'information a ouvert la voie de l'épidémiologie connectée ou épidémiologie digitale. Elle permet l'accès et l'analyse épidémiologique de nouvelles sources de données issues d'Internet, particulièrement utiles dans le champ de la surveillance sanitaire. La surveillance digitale est désormais établie en complément de la surveillance traditionnelle qui s'appuie sur la participation active de professionnels de santé. Les sources de données sont variées : données de santé (dossiers médicaux électroniques, données hospitalières) ou liées à la santé (délivrance de médicaments), mais aussi données externes (moteurs de recherche, blogs, réseaux sociaux). Ces différentes sources sont presque toujours extraites de façon passive, la quantité d'informations recueillies est très importante et les analyses complexes. En parallèle, la surveillance participative s'est développée pour pallier les limites de cette surveillance passive. Elle permet un suivi épidémiologique classique, mais directement à partir de la population et en utilisant les atouts d'Internet. Depuis 2012 en France, la cohorte participative GrippeNet.fr permet la surveillance complémentaire des syndromes grippaux directement à partir de la population résidant sur le territoire métropolitain.

Dans cette thèse, on s'intéressera à la description du système de surveillance de la grippe de l'échelle internationale à l'échelle nationale, en détaillant l'émergence et l'utilisation de sources de données non traditionnelles. Un focus sera effectué sur le rôle de ces sources dans le cadre de la surveillance de la grippe, et sur leurs avantages et limites. Le système européen de surveillance participative InfluenzaNet sera présenté, et celui de GrippeNet.fr détaillé, afin d'estimer leur valeur ajoutée dans la surveillance des syndromes grippaux.

Nous présenterons un premier travail dont l'objectif était d'étudier les nouveaux enjeux rencontrés avec le système participatif GrippeNet.fr. Les participants issus de la population générale sont-ils aptes à retranscrire leur état de santé ? A quel point le système est-il flexible pour permettre le suivi des maladies grippales ? Comment les participants déclarent-ils leurs symptômes et comment leur participation dynamique impacte-t-elle le signal estimé ? Grâce à cette étude, un cadre standardisé d'analyse du signal a été mis en place pour estimer les incidences des maladies grippales directement à partir de la population générale.

Un second travail a été mené pour analyser et valider les données épidémiologiques recueillies auprès des individus. Ces données sont-elles fiables pour mener des projets de recherche en population générale ? A titre d'exemple, nous avons étudié les facteurs associés aux maladies grippales à partir de GrippeNet.fr. La plupart était retrouvés dans la littérature, mais la finesse des données collectées a permis d'en identifier des nouveaux.

Ces deux travaux valident l'utilisation de la cohorte participative GrippeNet.fr à l'échelle populationnelle et individuelle, pour la mise en place d'une surveillance de la grippe et de projets de recherche en population générale.

Mots-clés : grippe humaine ; maladies infectieuses ; sciences participatives ; cohortes ; surveillance épidémiologique ; incidence ; facteurs épidémiologiques

Thesis summary

Participatory surveillance and influenza epidemiology among the general population

The emergence of the Internet and the speed of information exchange has paved the way for digital epidemiology. It allows the epidemiological analysis of new data sources from the Internet, which are particularly useful in the field of health surveillance. Digital surveillance is now established as a complement to traditional surveillance based on the active participation of health professionals. Data sources are diverse and are linked to: health data (electronic medical records, hospital data) or health-related data (drug delivery), but also external data (search engines, blogs, social networks). These different sources are almost always passively extracted, the amount of information collected is very large and the analyses complex. Participatory surveillance has been developed to overcome the limitations of this passive monitoring. It allows classical epidemiological monitoring, but directly from the population and using the advantages of the Internet. Since 2012 in France, the GrippeNet.fr participative cohort has been providing complementary surveillance of influenza illnesses directly from the population living in France.

This thesis will focus on describing the influenza surveillance system from the international to the national level, detailing the emergence and use of non-traditional data sources. A focus will be placed on the role of these sources in influenza surveillance, and their advantages and limitations. The European participatory surveillance system InfluenzaNet will be presented, and that of GrippeNet.fr detailed, in order to estimate the added value of this type of surveillance.

We will present a first work whose objective was to study the new challenges encountered with the participative system GrippeNet.fr. Are participants from the general population able to record their health status? How flexible is the system to monitor influenza diseases? How do participants report their symptoms and how does their dynamic participation impact the estimated signal? As a result of this study, a standardized signal analysis framework was implemented to estimate the incidence of influenza diseases directly from the general population.

A second study was carried out to analyse and validate the epidemiological data obtained through participative surveillance. Is this data reliable for conducting research projects in the general population? As a case study, we studied the factors associated with influenza diseases from GrippeNet.fr. Most of them were found in the literature, but the fine scale data collected allowed the identification of new ones.

These two studies validate the use of the participative cohort GrippeNet.fr at a population and individual level, for the implementation of influenza surveillance and of targeted research projects from the general population.

Keywords: human influenza; infectious disease; crowdsourcing; cohort studies; epidemiological monitoring; incidence; epidemiologic factors

Production scientifique réalisée dans cette thèse

Articles

Guerrisi C, Turbelin C, Souty C, Poletto C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, Colizza V. **The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France.** Euro Surveill. 2018. 23(25).

Guerrisi C, Ecollan M, Souty C, Rossignol L, Turbelin C, Debin M, Goronflot T, Boëlle PY, Hanslik T, Colizza V, Blanchon T. **Factors associated with influenza-like-illness: a crowdsourced cohort study from 2012/13 to 2017/18.** BMC Public Health. 2019. 19(1):879.

Communications orales

Guerrisi C. **GrippeNet.fr: une cohorte en population générale via Internet depuis 5 saisons (2012-2016).** Summer School of infectious diseases, 15-17 novembre 2016, Institut d'études scientifiques, Cargèse, France.

Guerrisi C. **GrippeNet.fr : Incidences, recours aux soins, facteurs de risque.** Séminaire scientifique sur la grippe - Santé publique France, 19 octobre 2018, Saint-Maurice, France.

Guerrisi C. **GrippeNet.fr: from population to individual levels.** Global Health Forum, April 10th-12th 2018, Geneva, Switzerland.

Production scientifique associée au travail de thèse

Articles

Guerrisi C, Turbelin C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, Perrotta D, Paolotti D, Smallenburg R, Koppeschaar C, Franco AO, Mexia R, Edmunds WJ, Sile B, Pebody R, van Straten E, Meloni S, Moreno Y, Duggan J, Kjelsø C, Colizza V. **Participatory Syndromic Surveillance of Influenza in Europe.** J Infect Dis. 2016. 214(suppl 4):S386-S392.

Geneviève LD, Martani A, Wangmo T, Paolotti D, Koppeschaar C, Kjelsø C, Guerrisi C, Hirsch M, Woolley-Meza O, Lukowicz P, Flahault A, Elger BS. **Participatory Disease Surveillance Systems: Ethical Framework.** J Med Internet Res. 2019. 21(5):e12273.

Kalimeri K, Delfino M, Cattuto C, Perrotta D, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds J, Obi C, Pebody R, Franco AO, Moreno Y, Meloni S, Koppeschaar C, Kjelsø C, Mexia R, Paolotti D. **Unsupervised extraction of epidemic syndromes from participatory influenza surveillance self-reported symptoms.** PLoS Comput Biol. 2019. 15(4):e1006173.

Mathieu P, Gautier A, Raude J, Goronflot T, Launay T, Debin M, Guerrisi C, Turbelin C, Hanslik T, Jestin C, Colizza V, Blanchon T, Rossignol L. **Population perception of mandatory childhood vaccination programme before its implementation, France, 2017.** Euro Surveill. 2019. 24(25).

Ung A, Baidjoe AY, Van Cauteren D, Fawal N, Fabre L, Guerrisi C, Danis K, Morand A, Donguy MP, Lucas E, Rossignol L, Lefèvre S, Vignaud ML, Cadel-Six S, Lailler R, Jourdan-Da Silva N, Le Hello S. **Disentangling a complex nationwide Salmonella Dublin outbreak associated with raw-milk cheese consumption, France, 2015 to 2016.** Euro Surveill. 2019. 24(3).

Ariza M, Guerrisi C, Souty C, Rossignol L, Turbelin C, Hanslik T, Colizza V, Blanchon T. **Healthcare-seeking behaviour in case of influenza-like illness in the French general population and factors**

associated with a GP consultation: an observational prospective study. *BJGP Open*. 2018. 1(4):bjgpopen17X101253.

Koppeschaar CE, Colizza V, **Guerrisi C**, Turbelin C, Duggan J, Edmunds WJ, Kjelsø C, Mexia R, Moreno Y, Meloni S, Paolotti D, Perrotta D, van Straten E, Franco AO. **Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe**. *JMIR Public Health Surveill*. 2017. 3(3):e58.

Loubet P, **Guerrisi C**, Turbelin C, Blondel B, Launay O, Bardou M, Blanchon T, Bonmarin I, Goffinet F, Ancel PY, Colizza V, Hanslik T, Kernéis S. **First nationwide web-based surveillance system for influenza-like illness in pregnant women: participation and representativeness of the French G-GrippeNet cohort**. *BMC Public Health*. 2016. 16(1):253.

Loubet P, **Guerrisi C**, Turbelin C, Blondel B, Launay O, Bardou M, Goffinet F, Colizza V, Hanslik T, Kernéis S, GNET study group. **Influenza during pregnancy: Incidence, vaccination coverage and attitudes toward vaccination in the French web-based cohort G-GrippeNet**. *Vaccine*. 2016. 34(20):2390-6.

Table des matières

LISTE DES ILLUSTRATIONS ET DES TABLEAUX.....	11
ABRÉVIATIONS ET ACRONYMES.....	12
INTRODUCTION	13
CHAPITRE 1. SURVEILLANCE EPIDEMIOLOGIQUE TRADITIONNELLE DE LA GRIPPE.....	17
1.1. Définitions et objectifs.....	17
1.2. L'épidémiologie de la grippe.....	18
1.2.1. Grippe ou syndrome grippal ?	19
1.2.2. Surveillance dans le monde	20
1.2.3. Surveillance en Europe	21
1.2.4. Le système de surveillance de la grippe en France métropolitaine	22
1.2.5. Analyse des données de morbidité en soins primaires, l'exemple du réseau Sentinelles	28
1.2.6. Avantages et limites de la surveillance traditionnelle de la grippe	29
CHAPITRE 2. LA SURVEILLANCE ET L'EPIDEMIOLOGIE DIGITALE APPLIQUEE A LA GRIPPE.....	33
2.1. Sources de données et applications.....	34
2.1.1. Description des principaux systèmes de surveillance utilisant Internet comme source d'information..	35
2.1.2. Les moteurs de recherche comme source d'information	37
2.1.3. Les réseaux sociaux comme source d'information.....	39
2.2. Avantages et limites.....	40
CHAPITRE 3. SCIENCES PARTICIPATIVES ET SURVEILLANCE DES MALADIES INFECTIEUSES	43
3.1. Principes de la surveillance participative et applications.....	43
3.2. InfluenzaNet ou la surveillance participative des syndromes grippaux en Europe.....	46
3.2.1. Collecte et gestion des données	47
3.2.2. Représentativité des données collectées.....	48
3.2.3. Incidence des syndromes grippaux estimée par les systèmes participatifs en Europe	49
3.2.4. Autres travaux de recherche menés par le biais de la surveillance participative des syndromes grippaux en Europe	51
3.3. La surveillance participative en France avec GrippeNet.fr.....	52
3.3.1. Recrutement, participation et communication	53
3.3.2. GrippeNet.fr pour mieux comprendre la grippe	57
3.4. Avantages et limites de la surveillance participative	58
CHAPITRE 4. L'INTERET POTENTIEL DES SYSTEMES DE SURVEILLANCE PARTICIPATIFS POUR COMPLETER LES RESEAUX SENTINELLES DE LA GRIPPE : LE CAS DE LA FRANCE.....	61
4.1. Problématiques liées à l'estimation des incidences des syndromes grippaux	61
4.2. Résumé de l'article	63
4.3. Article : The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France	64

4.4. Perspectives	76
CHAPITRE 5. FACTEURS ASSOCIES AUX SYNDROMES GRIPPAUX A PARTIR D'UNE COHORTE PARTICIPATIVE EN POPULATION GENERALE DE 2012/2013 A 2017/2018	79
5.1. Résumé de l'article	80
5.2. Article : Factors associated with influenza-like-illness : a crowdsourced cohort study from 2012/13 to 2017/18	81
5.3. Perspectives	92
CHAPITRE 6. APPLICATION DE GRIPPENET.FR AU-DELA DE LA GRIPPE.....	93
6.1. GrippeNet.fr pour le suivi d'autres maladies infectieuses.....	93
6.2. GrippeNet.fr au-delà des maladies infectieuses.....	95
CONCLUSION GENERALE	99
REFERENCES.....	103
ANNEXE 1. QUESTIONNAIRE D'INCLUSION	113
ANNEXE 2. QUESTIONNAIRE HEBDOMADAIRE DE SYMPTOMES	118

Liste des illustrations et des tableaux

Figure 1. Le Global Influenza Surveillance and Response System de l'Organisation mondiale de la Santé et la localisation de ses 152 centres nationaux	21
Figure 2. Le système de surveillance de la grippe en France	23
Figure 3. Transmission des données dans le cadre du dispositif de surveillance de la grippe, à partir des systèmes de surveillance du réseau Sentinelles et SurSaUD®	24
Figure 4. Proportions de la population évaluées par les systèmes de surveillance traditionnels et par Internet pendant une épidémie de grippe dans les pays à revenu élevé et faible	34
Tableau 1. Caractéristiques de trois systèmes traitant les sources de données informelles basées sur des événements.....	36
Figure 5. Localisation des modérateurs et abonnés et maladies recensées selon les régions du monde entre janvier 2016 et décembre 2016 par le système ProMED-mail.....	36
Figure 6. Utilisation d'Internet par pays.....	39
Tableau 2. Principaux systèmes de surveillance participatifs à travers le monde, pour le suivi de maladies infectieuses.....	45
Figure 7. Représentation de la collecte des données entre octobre/novembre et avril/mai au sein de la plateforme InfluenzaNet.....	48
Figure 8. Participation à GrippeNet.fr.	54
Figure 9. Satisfaction des participants lors de la saison 2015/2016.....	55
Figure 10. Communications réalisées pour la diffusion des résultats issus de l'analyse des données collectées par le biais de GrippeNet.fr.....	56
Figure 11. Types et nombre d'études menées avec la cohorte GrippeNet.fr	93
Figure 12. Taux de gastroentérites et d'allergies respiratoires parmi les participants de GrippeNet.fr lors de la saison 2018/2019.	94
Figure 13. Aperçu des caractéristiques des systèmes de surveillance des maladies infectieuses	102

Abréviations et acronymes

CDC : Centers for Disease Control and Prevention
CépiDC : Centre d'épidémiologie sur les causes médicales de décès
CIM-10 : Classification internationale des maladies (10^{ème} révision)
Cire : Cellule d'intervention en région (SpF en région)
CNR : Centre national de référence
DGAL : Direction générale de l'alimentation
ECDC : European Centre for Disease Prevention and Control
Ehpad : Etablissement d'hébergement pour personnes âgées dépendantes
EISN : European Influenza Surveillance Network
Eprus : Etablissement de préparation et de réponse aux urgences sanitaires
EuroMomo : European monitoring of excess mortality for public health action
GFT : Google Flu Trends
GISN : Global Influenza Surveillance Network
GISRS : Global Influenza Surveillance and Response System
GPHIN : Global Public Health Intelligence Network
GROG : Groupes Régionaux d'Observation de la Grippe
ILI : Influenza-like illness
Inpes : Institut national de prévention et d'éducation à la santé
Insee : Institut national de la statistique et des études économiques
Inserm : Institut national de la santé et de la recherche médicale
InVS : Institut national de veille sanitaire
IRA : Infection respiratoire aiguë
MASS : Module d'Analyse des données SurSaUD et Sentinelles
MDO : Maladie à Déclaration Obligatoire
MEM : Moving Epidemic Method
MGS : Médecins généralistes Sentinelles
OMS : Organisation Mondiale de la Santé
OSCOUR : organisation de la surveillance coordonnée des urgences
RENAL : Réseau national des laboratoires hospitaliers
RGPD : Règlement général sur la protection des données
RSI : Règlement Sanitaire International
SAMU : Service d'Aide Médicale Urgente
SMF : SOS Médecins France
SpF : Santé publique France
SRAS : syndrome respiratoire aigu sévère
SurSaUD : Surveillance Sanitaire des Urgences et des Décès
TESSy : The European Surveillance System
USA : United States of America
VPP : Valeur prédictive positive
VRS : Virus syncytial respiratoire

Introduction

La surveillance épidémiologique repose sur des systèmes qui collectent, analysent et diffusent des informations utiles au suivi et au maintien de la santé publique. Les indicateurs de santé servent à décrire la situation épidémiologique dans le temps et l'espace, alerter les acteurs de santé publique en cas de signal anormal, et évaluer les réponses apportées [1, 2]. La surveillance épidémiologique est surtout appliquée au suivi des maladies infectieuses, face auxquelles elle constitue le principal moyen de lutte [3, 4].

Les systèmes de surveillance collectent des données issues de sources multiples et diverses, mais leurs rôles restent similaires : traduire au mieux la situation de l'indicateur de santé donné grâce aux outils épidémiologiques disponibles. Un système de surveillance se doit d'être le plus efficace et efficient possible. Dans ce but, le choix des critères de qualité à prendre en compte dépend de l'objectif de la surveillance, mais ceux-ci reposent classiquement sur les caractéristiques suivantes : pertinence, flexibilité, acceptabilité, simplicité, sensibilité, valeur prédictive positive, réactivité, représentativité [1, 5]. Les systèmes de surveillance ayant pour objectif d'alerter les autorités sanitaires doivent être réactifs, donc simples et acceptables, mais aussi représentatifs voire sensibles [1, 5].

La surveillance épidémiologique traditionnelle s'appuie sur l'intervention volontaire de professionnels de santé (déclaration obligatoire, réseaux de type sentinelle), qui permettent une collecte spécifique de données, avec le recueil d'informations fiables et précises. Elle tend à s'ouvrir à d'autres sources et moyens de collecte pour combler des lacunes de réactivité, flexibilité, sensibilité et représentativité, et cela d'autant plus depuis la pandémie grippale A(H1N1) de 2009 et l'émergence de nouvelles maladies.

Ainsi, depuis les années 2000 et grâce à l'essor d'Internet, la surveillance s'est développée avec l'extraction directe des données à partir de sources de santé (données administratives, médicales comme les données hospitalières, issues des ventes de médicaments, etc.) ou d'autres sources, presque toujours non conçues pour le suivi épidémiologique (moteurs de recherche, blogs et microblogs, réseaux sociaux). La population générale devient elle aussi, a fortiori, actrice de cette surveillance : (i) soit passivement, par les traces laissées lors de recherches ou d'activités effectuées sur Internet ; (ii) soit activement, en prenant part de façon volontaire à la veille épidémiologique via la surveillance participative. L'avantage principal de l'extraction de données réside dans la collecte passive et automatisée d'informations de santé directement issues du système de soins ou d'Internet. Les données issues d'Internet permettent d'atteindre directement la population générale. Cependant, la survenue d'un phénomène de santé peut générer une quantité importante d'informations non forcément pertinentes. Les données collectées peuvent donc mesurer d'autres phénomènes que celui recherché [6]. Pour pallier ce

problème, certains systèmes ont été calibrés sur les données existantes issues du système de soins, limitant ainsi leur intérêt [7].

Dans ce contexte, le développement de la surveillance participative représente une alternative intéressante. D'un côté, elle permet de collecter des données directement à partir de la population générale, y compris au niveau de la population qui n'a pas recours au système de soins. De l'autre, elle bénéficie de la réactivité d'Internet et est indépendante des autres sources de données.

Différents systèmes de surveillance participatifs ont vu le jour ces dernières années. En France, la cohorte participative GrippeNet.fr a été initiée en 2012 pour suivre l'évolution de la grippe, directement à partir de la population résidant sur le territoire métropolitain. L'objectif de cette thèse sera alors de valider ce système participatif comme nouvelle source de données en population générale, en complément de la surveillance existante. Afin de justifier son utilisation dans le dispositif de surveillance de la grippe français, il est important de comprendre le fonctionnement propre du système, mais aussi au regard des autres systèmes en place. Cela passe par l'étude conjointe de (i) la validité du signal perçu, mesuré à partir d'une population non professionnelle, aussi bien au niveau populationnel qu'individuel ; (ii) l'évaluation de la flexibilité du système avec la reconstruction d'épisodes grippaux à partir de simples symptômes ; (iii) la prise en compte d'une participation hétérogène qui peut impacter l'estimation faite du signal.

Nous présenterons un premier travail qui nous a permis de valider les signaux d'incidence émis par cette cohorte, en évaluant la participation dynamique et la flexibilité du système. Cette étude propose une approche standardisée d'estimation des incidences propre à ce nouveau système de surveillance, estime la tendance et l'amplitude du signal, et évalue le bénéfice de GrippeNet.fr dans la surveillance de la grippe. Le second travail réalisé s'est concentré à valider plus finement ce système participatif par le biais des données épidémiologiques collectées. Ces analyses complémentaires portent sur la mise en évidence des facteurs associés aux épisodes grippaux en population, en confirmant et complétant ceux observés via le système de soins.

Dans le **Chapitre 1**, nous décrivons le système de surveillance traditionnelle de la grippe dans le monde et en Europe, avec une attention particulière portée au dispositif de surveillance français, qui s'appuie sur la déclaration de cas observés par des professionnels de santé, et sur l'extraction de données de santé issues du système de soins. Le **Chapitre 2** présentera la surveillance épidémiologique digitale, avec les différents systèmes développés en prenant l'exemple du suivi de la grippe. Dans le **Chapitre 3**, nous détaillerons spécifiquement le système de surveillance participatif européen InfluenzaNet. Nous présenterons son fonctionnement avec la collecte, la gestion et la représentativité des données collectées, mais aussi les travaux menés autour de l'estimation du signal grippal suivi, ainsi que les autres projets réalisés : estimation de l'efficacité vaccinale, recours aux soins en cas de syndrome grippal, déterminants de l'infection par un syndrome grippal. Nous nous focaliserons ensuite sur la cohorte GrippeNet.fr, en décrivant la communication faite et l'effort réalisé sur le maintien de la participation, qui représente un point crucial pour la pérennité du projet. Nous exposerons les études menées spécifiquement en

France pour mieux comprendre la grippe, comme le suivi de groupes de populations spécifiques ou l'estimation de l'opinion quant à la vaccination antigrippale. Le **Chapitre 4** exposera notre travail réalisé pour valider cette cohorte participative à un niveau populationnel. Nous proposerons une méthode d'estimation standardisée des incidences des maladies grippales à partir de ce type de surveillance participative, en prenant en compte les enjeux internes à cette nouvelle surveillance. Dans le **Chapitre 5** nous détaillerons un second travail, réalisé cette fois pour valider notre nouveau système de surveillance à un niveau individuel. Il présentera les facteurs associés à une infection par syndrome grippal dans, et à partir de la population générale. Enfin, le **Chapitre 6** portera sur les autres projets réalisés grâce à la cohorte participative GrippeNet.fr au-delà de la surveillance de la grippe et des maladies infectieuses. Nous détaillerons les études réalisées pour suivre l'évolution des gastroentérites hivernales, investiguer des épidémies d'origine alimentaire, ou évaluer des mesures utiles au développement d'actions de prévention et de promotion de la santé.

Chapitre 1.

Surveillance épidémiologique traditionnelle de la grippe

1.1. Définitions et objectifs

Depuis les années 1950 et la création des *Centers for Disease Controls* américains, la surveillance épidémiologique a été définie comme un « processus systématique de collecte, d'analyse et d'interprétation de données sur des événements de santé spécifiques importants pour la planification, la mise en œuvre, et l'évaluation des pratiques en santé publique » [1]. Sur ce principe, la surveillance épidémiologique s'organise en fonction du système de soins établi dans le pays. En France et dans les pays occidentaux, différents types de surveillance sont alors possibles, à partir de sources de données variées (administratives, médicales, biologiques, environnementales, ou issues de cohortes, d'enquêtes). Le système est ensuite tourné vers l'alerte ou vers l'évaluation selon les objectifs de la surveillance et les caractéristiques de l'événement surveillé. Un système d'alerte est adapté à la détection de phénomènes aigus comme le suivi d'épidémies ou de l'émergence de pathologies, nécessitant la mise en place d'une action de santé publique rapide voire urgente. Les systèmes tournés vers l'évaluation viseront plutôt à étudier les tendances des maladies ou encore l'impact d'une politique de santé publique sur un espace spatio-temporel défini. Pour jouer son rôle d'alerte ou d'évaluation, un système de surveillance épidémiologique pourra fonctionner par le biais (i) d'obligations réglementaires, comme le dispositif des maladies à déclaration obligatoire (MDO), ou par (ii) des réseaux de professionnels hospitaliers ou libéraux, en exercice particulier ou fédérés au sein d'associations [1, 8].

Les systèmes mis en place sont souvent spécifiques à une pathologie donnée, cependant d'autres approches complémentaires peuvent être utilisées afin de répondre au mieux à la détection de menaces sanitaires diverses (i.e. phénomène environnemental, maladie émergente) en temps quasi-réel. On retrouve alors la surveillance syndromique, définie par le *Center for Diseases Control and Prevention* (CDC) comme « une approche, dans laquelle les intervenants sont assistés par des procédures d'enregistrement automatiques des données, qui permettent leur mise à disposition pour le suivi et l'analyse épidémiologique en temps réel ou quasi-réel. Cela afin de détecter des événements habituels ou inhabituels plus tôt qu'il n'aurait été possible de le faire sur la base des méthodes traditionnelles de surveillance » [9]. SurSaUD® (Surveillance sanitaire

des urgences et des décès) est un exemple français de surveillance syndromique, mis en place et coordonné par Santé publique France (SpF). Depuis 2004, ce système permet la collecte quotidienne de données non spécifiques provenant de services d'urgences, d'associations de médecine d'urgence et de communes pour les données de mortalité. Au-delà de l'approche traditionnelle et syndromique, d'autres sources d'information peuvent être utilisées pour des fonctions d'alerte précoce. On retrouve notamment la surveillance environnementale et écologique (i.e. données climatiques, qualité de l'eau, etc.) et l'information comportementale associée à la santé (i.e. absentéisme, ventes de médicaments, activités sur internet, etc.). L'objectif d'une surveillance multi-sources est donc d'être sensible pour détecter et répondre rapidement aux signaux et alertes venus des différentes sources formelles ou informelles [10].

La mise en place de dispositifs de surveillance croisant différentes sources d'information sur différents niveaux permet d'avoir une vision globale de la situation du phénomène de santé, et de pouvoir suivre au mieux son évolution spatio-temporelle. En effet, l'objectif principal de la veille sanitaire est de répondre au mieux à la détection en temps quasi-réel de l'occurrence de menaces sanitaires pour en informer les décideurs. La description et la compréhension globale d'un événement de santé permet l'application d'actions de santé publique adaptées et ciblées pour lutter efficacement contre le phénomène en cause qui impacte la santé de la population.

Dans les paragraphes suivants, nous allons décrire le dispositif de surveillance de la grippe - en particulier la morbidité inhérente - en se focalisant sur le système français. En effet, ce dispositif a su tirer profit de deux événements sanitaires en France, la canicule de 2003 et la grippe pandémique A(H1N1) en 2009. Après 2003, le dispositif s'est alors tourné vers la mise en place d'une approche syndromique en plus de l'approche spécifique existante. Après 2009, le dispositif s'est doté d'un système de surveillance des cas graves et groupés. Aujourd'hui, le dispositif de surveillance de la grippe est l'un des plus complets. Il permet un suivi à différents niveaux et à partir de différentes sources de données, que nous allons détailler.

1.2. L'épidémiologie de la grippe

La grippe est une infection virale aiguë qui se transmet facilement d'une personne à l'autre. Dans les régions tempérées du globe, des épidémies de grippe saisonnière ont lieu chaque année au cours des mois d'hiver (entre novembre et mars dans l'hémisphère Nord, et entre mai et septembre dans l'hémisphère Sud). Dans certains pays tropicaux, on ne retrouve pas cette saisonnalité caractéristique des pays tempérés, les virus de la grippe y circulent toute l'année, avec un ou plusieurs pics à la saison des pluies [11, 12].

Un épisode grippal se caractérise par l'apparition brutale d'une fièvre accompagnée de signes respiratoires et de myalgies. La majorité des personnes qui contractent une grippe saisonnière se rétablissent sans l'aide d'un professionnel de santé. Cependant, l'infection peut entraîner des complications (dues au virus lui-même ou à des surinfections bactériennes), voire être mortelle,

en particulier chez les jeunes enfants, les personnes âgées (65+), les sujets fragiles (femmes enceintes), ou ceux atteints de maladies chroniques concomitantes (obésité, affections broncho-pulmonaires chroniques, pathologies cardiaques graves, néphropathies et hépatopathies chroniques graves, diabète, immunodépression, etc.). Chaque année selon les souches virales circulantes, on estime que la grippe est à l'origine de 3 à 5 millions de pathologies graves et de 290 000 à 650 000 décès à travers le monde [12, 13]. En France, une épidémie de grippe saisonnière dure en moyenne 9 semaines, et 4,1% (IC95% [3,5 ; 4,7]) de la population est atteinte [14]. En plus de l'impact sanitaire, une épidémie de grippe va donc avoir un impact économique. Dans les pays développés, celle-ci peut ainsi causer un fort absentéisme parmi les travailleurs et entraîner des pertes de productivité. En France, une épidémie saisonnière coûterait 2,6 milliards de dollars (US) à la société, dont 0,3 milliards de coûts sanitaires directs et 2,3 milliards de coûts indirects dus à la perte de productivité, avec 2,9 ($\pm 2,5$) jours de travail perdus par épidémie et par personne [15, 16]. Les effets des épidémies de grippe dans les pays en développement sont plutôt mal connus, faute de données pertinentes [17].

Si toutes les maladies infectieuses ne font pas l'objet d'une surveillance, la grippe fait partie des maladies pour lesquelles une surveillance est jugée indispensable [18, 19]. En effet, plusieurs facteurs au niveau national et international en font une maladie infectieuse prioritaire : fréquence élevée, transmissibilité et risque d'épidémie, gravité, coût social élevé et existence de mesures de contrôle, de prévention ou de prise en charge efficaces [20].

1.2.1. Grippe ou syndrome grippal ?

Le suivi de la grippe s'effectue principalement à partir d'un diagnostic clinique effectué par un professionnel de santé. Ce diagnostic repose sur une définition de cas propre au système de surveillance auquel est rattaché le professionnel. Différentes définitions coexistent alors, aussi bien à l'échelle internationale, qu'à l'échelle nationale, ce qui peut compliquer la coordination du système de surveillance [21-23]. Les différentes définitions de cas s'appuient sur l'inclusion de symptômes systémiques (fièvre, frissons, fatigue, myalgies, mal de tête) et respiratoires (toux, mal de gorge, dysphagie, rhinorrhée ou congestion nasale). L'association de plusieurs symptômes permet un suivi clinique de la grippe, à travers deux indicateurs principalement utilisés dans les systèmes de surveillance : le syndrome grippal et l'infection respiratoire aiguë (IRA). *Aguilera et al.* ont étudié les différentes définitions de cas utilisées à travers des pays d'Europe de l'ouest. Sur les 21 réseaux de surveillance répondants, plus des deux tiers (n=15) utilisaient une définition de syndrome grippal, le tiers restant se partageait entre l'utilisation d'une définition d'IRA (n=3) ou des deux indicateurs (n=3) [21]. D'autres travaux ont été menés pour savoir quels sont les symptômes qui prédisent le mieux la grippe. Une fièvre supérieure ou égale à 39°C, de la toux et des maux de gorges étaient souvent associés à la présence du virus grippal [22, 23].

Dans l'ensemble, les définitions de syndrome grippal sont plus spécifiques que les définitions d'IRA. L'utilisation d'une définition plus spécifique ou plus sensible va dépendre du système de

surveillance recherché. Un système qui aura pour seul objectif de surveiller la grippe aura tout intérêt à utiliser une définition la plus spécifique possible, pour ne capter majoritairement que les malades liés aux virus grippaux. Au contraire, un système de surveillance sensible permettra d'intégrer des malades liés aux virus grippaux ou à d'autres virus respiratoires, comme le virus syncytial respiratoire (VRS), le rhinovirus ou encore le métapneumovirus [24]. *Jiang et al.* ont montré que l'utilisation de définitions d'IRA était inappropriée pour suivre la grippe [22]. Cependant, les définitions d'IRA s'avèrent utiles pour le suivi de la bronchiolite chez les nourrissons ou des IRA chez les personnes âgées, souvent dues au VRS et qui représentent un fardeau important pour les deux populations [25, 26]. Quoi qu'il en soit, seule une analyse virologique effectuée après un prélèvement naso-pharyngé, pourra confirmer ou non la présence de virus grippaux ou d'autres virus respiratoires responsables de la maladie.

1.2.2. Surveillance dans le monde

Depuis 1952 au niveau mondial, l'Organisation mondiale de la Santé (OMS) surveille la grippe, initialement via le *Global Influenza Surveillance Network* (GISN) puis par le biais du *Global Influenza Surveillance and Response System* (GISRS). Ce dernier joue un rôle central dans la détection et la surveillance de la grippe, à la fois sur le versant épidémiologique et virologique. Il permet l'identification des souches virales circulantes et surveille leur évolution, détecte et évalue le potentiel pandémique de nouveaux virus grippaux, suit la progression de l'épidémie, et évalue ses conséquences (cas graves, mortalité). Le GISRS est doté d'un réseau de laboratoires qui analysent les virus grippaux et en suivent l'évolution. A partir des résultats obtenus, l'OMS établit les recommandations pour la composition du vaccin antigrippal saisonnier dans l'hémisphère Nord et dans l'hémisphère Sud. Cinq centres mondiaux, situés à Atlanta, Londres, Melbourne, Pékin et Tokyo, collectent les informations provenant de 152 centres nationaux (Figure 1) [27]. Si cet important système de surveillance permet au monde d'améliorer sa capacité à lutter contre les épidémies et pandémies de grippe, un renforcement du système en place, par l'amélioration de l'intégration des données et par des collaborations institutionnelles et intersectorielles, nationales et internationales, reste nécessaire pour optimiser la gestion des risques. Dans ce sens depuis 2005, le règlement sanitaire international (RSI) fait obligation aux Etats Membres de l'OMS de développer des capacités essentielles minimales en matière de surveillance, de rapport, de notification, de vérification, de riposte et d'activités de collaboration pour les événements de santé publique de portée internationale [28]. C'est l'instrument juridique international mis en place pour protéger tous les pays contre une propagation mondiale des maladies. La pandémie grippale de 2009 a été la première mise à l'épreuve véritable du fonctionnement et de l'application de ce règlement. Le RSI a contribué à ce que l'ensemble du monde soit mieux préparé aux urgences de santé publique.

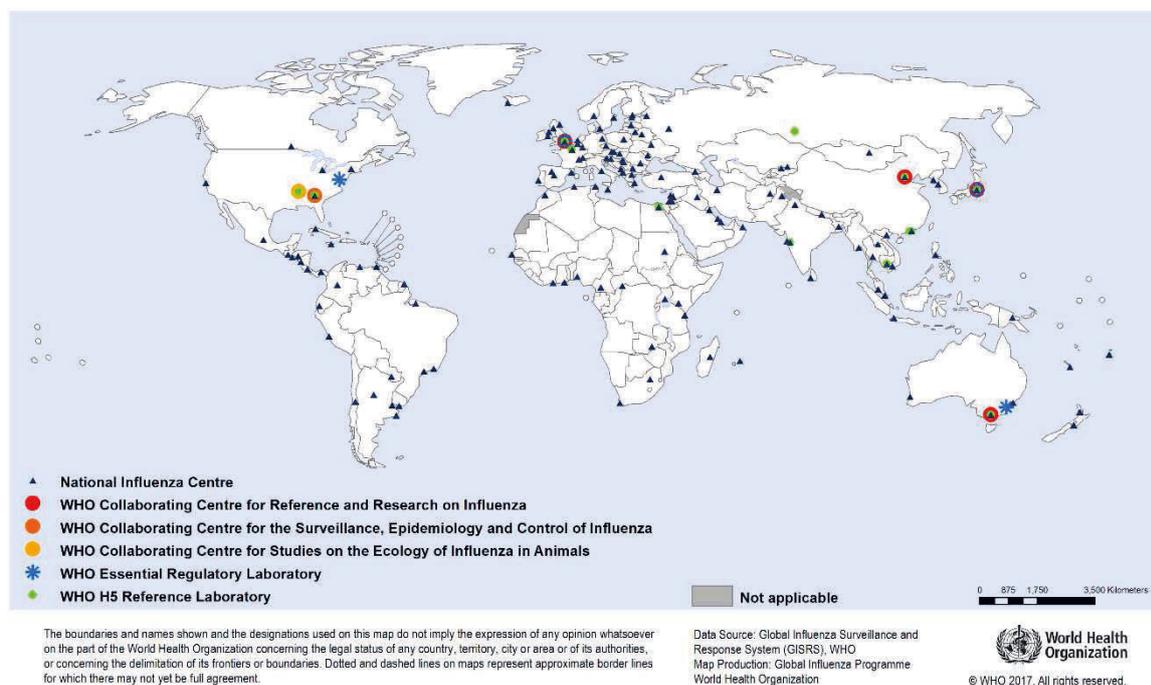


Figure 1. Le Global Influenza Surveillance and Response System de l’Organisation mondiale de la Santé et la localisation de ses 152 centres nationaux. *Source : OMS, 2017*

1.2.3. Surveillance en Europe

La surveillance de la grippe est coordonnée par le bureau européen de l’OMS et le *European Centre for Disease Prevention and Control* (ECDC). La mission principale de cette surveillance européenne est de suivre les maladies bénignes et graves dues à la grippe, (i) en informant les programmes nationaux de prévention et de contrôle de la gravité et de l’impact des épidémies saisonnières, (ii) en identifiant les souches virales circulantes, (iii) et en détectant les virus grippaux émergents (zoonotique et pandémique).

L’ECDC, à travers le *European Influenza Surveillance Network* (EISN) est en charge de la collecte et de l’analyse des données de surveillance de la grippe fournies par les Centres nationaux de la grippe des 53 Etats Membres. Les données nationales concernant les syndromes grippaux et IRA sont rapportées et collectées chaque semaine à travers une plateforme électronique commune nommée TESSy (The European Surveillance System) au moment de la saison grippale (entre la semaine 40 de l’année en cours et la semaine 20 de l’année suivante). La collecte des données épidémiologiques et virologiques est basée sur différentes sources : soins primaires (source majoritairement collectée et réalisée à partir de réseaux de médecins généralistes surtout, mais aussi de pédiatres, voire de médecins scolaires), soins hospitaliers (services d’urgences avant tout et parfois services d’hospitalisation), et données non sentinelles (hôpitaux, écoles, établissements de soins de santé primaires, maisons de retraite ou autres établissements ne participant pas à la surveillance sentinelle) [29]. Les données européennes sont ensuite présentées dans un bulletin hebdomadaire conjoint ECDC – OMS/Europe, nommé *Flu News Europe*

(<https://flunewseurope.org/>), qui fournit des mises à jour sur l'évolution de la saison grippale et la propagation de la grippe, sur la gravité de l'épidémie, et sur la prévalence et les caractéristiques des virus circulants. Ces données sont à destination des professionnels de santé surtout, et sont partagées deux fois par semaine avec le siège de l'OMS afin d'avoir une mise à jour de la situation de la grippe dans le monde [30].

La collecte des données à l'échelle européenne représente un enjeu majeur dans le cadre de surveillance de la grippe. Les différences rencontrées entre les systèmes de surveillance nationaux des Etats Membres peuvent être diverses. Cela peut concerner la nature des indicateurs surveillés, les moyens de surveillance utilisés ou encore les acteurs de cette surveillance. Les critères utilisés pour définir un syndrome grippal peuvent aussi différer d'un pays à l'autre en se basant sur des définitions de cas différentes (comme la définition de l'ECDC, celle de l'OMS ou d'autres spécifiques à l'Etat Membre). L'ECDC promeut l'harmonisation des moyens de surveillance au moyen de cinq indicateurs qualitatifs déclinables en différents niveaux. L'objectif est que chaque pays puisse, au regard de ses données et de son historique, caractériser l'épidémie de grippe sur son territoire. Les indicateurs concernent l'intensité de l'activité grippale perçue (déclinée en cinq niveaux : niveau de base, faible, modéré, fort, très fort); la distribution géographique du virus dans le pays (aucune dispersion, sporadique, locale, régionale, nationale); les tendances de l'évolution de l'épidémie au regard de semaines passées (en augmentation, stable, en diminution); le virus grippal circulant de façon dominante et l'impact en termes de gravité (i.e. hospitalisations, avec un niveau de base, faible, modéré, fort, très fort). Le choix du niveau des différents critères peut se faire soit de façon qualitative par le biais des experts nationaux, soit par une évaluation semi-quantitative s'appuyant sur les données historiques.

Au-delà de la surveillance de la morbidité liée à la grippe, une surveillance européenne de la mortalité s'effectue via le projet européen EuroMomo (European monitoring of excess mortality for public health action). Depuis 2008, son objectif est l'homogénéisation de l'enregistrement en routine des données de mortalité toutes causes dans les pays européens. La part de la mortalité attribuable à la grippe représente une sous-partie de ce système, pour laquelle un modèle statistique a été développé [31].

1.2.4. Le système de surveillance de la grippe en France métropolitaine

Le dispositif de surveillance de la grippe en France s'effectue à différents niveaux : de la population (population générale et soins de premiers recours), large et sensible, à un niveau plus spécifique, comme la surveillance des cas graves (i.e. hospitalisations, urgences, services de réanimation) ou des causes médicales de décès (Figure 2). Ces différents niveaux de surveillance forment un dispositif national global avec des acteurs variés qui doivent interagir ensemble pour une bonne efficacité du système. La coordination nationale de cette surveillance est réalisée par Santé publique France (SpF), l'agence nationale de santé publique. SpF est un établissement public administratif sous tutelle du ministère de la Santé. L'agence a été créée par le décret

n°2016-523 du 27 avril 2016 et fait partie de la loi de modernisation du système de santé. Elle regroupe l'Institut national de prévention et d'éducation pour la santé (Inpes), l'Institut national de veille sanitaire (InVS) et l'Établissement de préparation et de réponse aux urgences sanitaires (Eprus). Son rôle dans la surveillance de la grippe est la détection du début de l'épidémie saisonnière et sa description (suivi spatio-temporel, étude de sa gravité et identification des personnes à risque de complications, estimation de son impact dans la communauté et dans les structures de soins), l'évaluation de la vaccination et des mesures d'hygiène, et l'identification des souches virales circulantes.

Dans cette section, nous détaillerons la surveillance traditionnelle qui s'appuie sur la surveillance effectuée par les professionnels de santé, et la surveillance syndromique basée sur les données du système de soins. Les surveillances digitales et participatives seront détaillées dans les Chapitres 2 et 3.

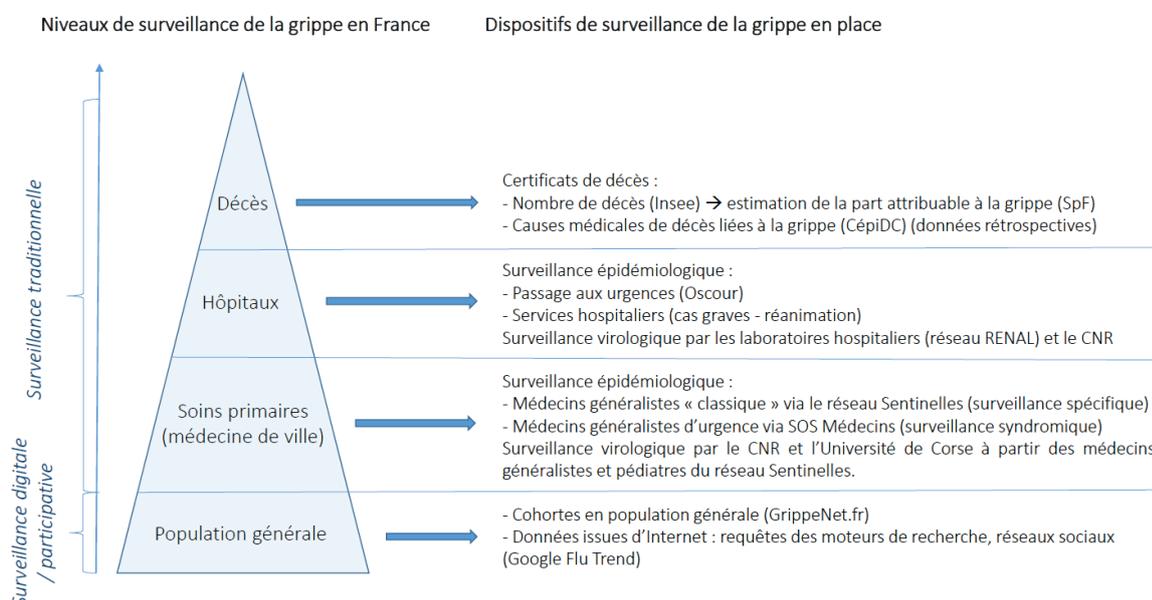


Figure 2. Le système de surveillance de la grippe en France

En France métropolitaine, le dispositif de surveillance de la grippe saisonnière est activé de la semaine 40 (début octobre) à la semaine 15 de l'année suivante (mi-avril). Chaque semaine des données épidémiologiques et virologiques sont collectées à travers les différentes structures (médecine de ville et hospitalière, collectivité, certificats de décès) et analysées. Le dispositif s'appuie sur (i) une surveillance de veille et d'alerte à partir de la participation volontaire de professionnels de santé en soins primaires, organisés en réseaux, qui effectuent un diagnostic clinique des syndromes grippaux, complété par une analyse virologique (cf. le réseau Sentinelles depuis 1984 ; le GROG entre 1984 et 2014) ; et sur (ii) une surveillance syndromique complémentaire mise en place depuis 2004, avec le système SurSaUD® (Surveillance Sanitaire des Urgences et des Décès) (Figure 3). Nous détaillerons ces deux surveillances dans les sections suivantes.

Afin de pouvoir regrouper et synthétiser l'information issue des réseaux de professionnels en soins primaires et des données de morbidité issues de la surveillance syndromique, SpF a développé en 2015 une application Web appelée MASS (Module d'Analyse des données SurSaUD® et Sentinelles). Cette application permet de combiner et d'analyser conjointement les données des trois sources de surveillance disponibles : (i) le réseau Sentinelles pour les incidences des syndromes grippaux en médecine générale ; (ii) SOS Médecins pour le nombre de consultations d'urgences en médecine générale ; et (iii) le réseau OSCOUR® pour le nombre de passages aux urgences pour syndrome grippal. Pour chacune des sources de données, MASS permet la génération d'une alarme épidémique selon différentes méthodes statistiques aux niveaux national et régional [32].

Finalement, le résultat couplé de l'analyse des données de morbidité et mortalité est retransmis à travers un bulletin épidémiologique à visée première des professionnels de santé, mais aussi du grand public.

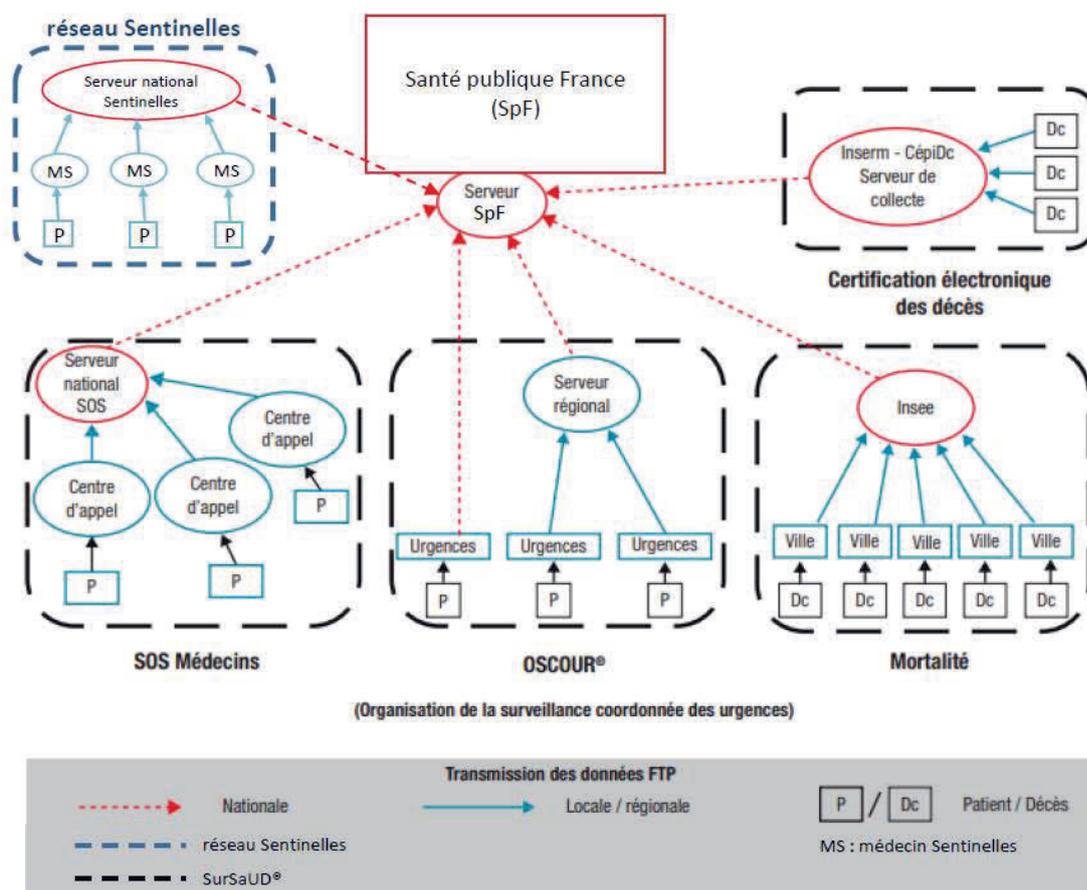


Figure 3. Transmission des données dans le cadre du dispositif de surveillance de la grippe, à partir des systèmes de surveillance du réseau Sentinelles et SurSaUD®. Source : Adaptation à partir de Caserio-Schönemann et al. [33]

1.2.4.1. Surveillance par les professionnels de santé en soins primaires

En France, et plus largement en Europe, le dispositif de surveillance en soins primaires repose principalement sur les médecins généralistes. En effet, ils représentent le premier contact du patient avec le système de soins qui ne nécessite pas forcément le recours à des soins hospitaliers [29, 34]. Les réseaux de surveillance en médecine générale sont un outil privilégié pour estimer l'état de santé d'une population et pour suivre son évolution au cours du temps, en particulier les épidémies [34]. Le choix des indicateurs de santé surveillés est défini sur la fréquence de la maladie dans la communauté et sur l'importance de sa morbidité. En ce sens, la surveillance de la grippe avec ses indicateurs cliniques (syndrome grippal et/ou IRA) et par ces réseaux est appropriée, puisque 2,4 millions de personnes sont touchées en moyenne lors d'une épidémie saisonnière en France [14].

Depuis 1984 et jusqu'en 2014, la surveillance de la grippe en soins primaires reposait sur deux réseaux de professionnels de santé : le GROG (Groupes Régionaux d'Observation de la Grippe) et le réseau Sentinelles. Le GROG, permettait le suivi clinique épidémiologique et virologique hebdomadaire des IRA à partir de médecins généralistes et de pédiatres libéraux [35]. Ce réseau de surveillance n'a finalement pas été maintenu et aujourd'hui seul le réseau Sentinelles perdure pour la surveillance de la grippe en soins primaires.

Le **réseau Sentinelles** permet depuis 1984 le recueil et l'analyse de données épidémiologiques en France métropolitaine, à partir de médecins généralistes et pédiatres libéraux [36, 37]. Il s'agit d'un réseau de recherche et de veille sanitaire en soins primaires, sous la tutelle de l'Inserm et de Sorbonne Université. La participation des médecins est volontaire et bénévole. Au 1^{er} janvier 2019, le réseau était composé de 1 334 médecins généralistes libéraux, soit 2,3% des médecins généralistes libéraux métropolitains, et de 116 pédiatres représentant 4,4% des pédiatres libéraux [38]. Chaque semaine, les médecins généralistes Sentinelles (MGS) déclarent les cas vus en consultation (ou l'absence de cas) pour différents indicateurs de santé. Au 1^{er} janvier 2019, 10 indicateurs de santé étaient suivis de façon continue par 480 MGS (soit 36% des 1 334 MGS inscrits). Sur les dix événements de santé suivis, neuf étaient infectieux : syndrome grippal, infection respiratoire aiguë (IRA), diarrhée aiguë, varicelle, zona, coqueluche, urétrites masculines, oreillons, maladie de Lyme, et un non-infectieux : les actes suicidaires. En plus de la déclaration des cas, les MGS ont la possibilité de décrire spécifiquement chacun des indicateurs (par exemple : âge, sexe, statut vaccinal s'il a lieu, maladies chroniques sous-jacentes, complications observées ou demandes d'hospitalisation), afin d'avoir des informations plus fines. Le recueil d'information est effectué directement sur un site dédié ou via un logiciel client, jSentinel [39]. La collecte des données permet entre autres : l'estimation de taux d'incidence hebdomadaires et annuels des différents indicateurs aux niveaux national et régional, et par classes d'âge [14, 40] ; la détection des épidémies saisonnières de grippe et gastroentérites [41, 42] et la prédiction de leur évolution [43] ; l'estimation de l'impact de mesures de santé publique sur ces indicateurs [44, 45]. Un retour d'information hebdomadaire est réalisé en parallèle par le

réseau Sentinelles pour trois indicateurs (varicelle, diarrhée aiguë, syndrome grippal), et un retour annuel est effectué pour les autres indicateurs dans un bilan annuel [38].

Pour le suivi des épidémies de grippe, le réseau Sentinelles collecte à la fois des données épidémiologiques cliniques (tout au long de l'année depuis 1984) et virologiques (de la semaine 40 à la semaine 15 de l'année suivante depuis 2014) sur le syndrome grippal par le biais des MGS. Les pédiatres ne participent qu'au suivi virologique. Cette surveillance s'effectue en collaboration avec SpF et le Centre national de référence (CNR) des virus des infections respiratoires (dont la grippe). L'objectif est de détecter et suivre la survenue des épidémies de grippe et de pouvoir les décrire ; estimer l'efficacité du vaccin antigrippal ; identifier les souches grippales, mais aussi les autres virus respiratoires circulants. D'ailleurs, un suivi clinique et virologique des IRA chez les personnes âgées de 65 ans et plus est venu compléter la surveillance du syndrome grippal en 2017, afin de capturer plus largement ces autres virus respiratoires dans cette tranche d'âge, pour mieux en comprendre l'impact. Cette nouvelle surveillance est toujours en phase pilote. Afin d'avoir une surveillance homogène entre les professionnels de santé, le réseau Sentinelles a défini le syndrome grippal comme une fièvre supérieure à 39°C, d'apparition brutale, accompagnée de myalgies et de signes respiratoires. Les IRA sont définies par l'apparition brutale d'au moins un des trois signes respiratoires parmi la toux, la dyspnée ou la rhinite, ne répondant pas à la définition d'un syndrome grippal et qui est due à une infection (selon le jugement du clinicien).

Au-delà de cette surveillance spécifique en soins primaires, une surveillance syndromique du syndrome grippal a lieu à partir des données de médecine générale d'urgence via les associations **SOS Médecins France** (SMF). Fondé en 1966, il s'agit d'un réseau de 61 associations d'urgentistes et de médecins généralistes libéraux réparties sur le territoire métropolitain. Ces associations sont présentes dans la plupart des grands centres urbains et leur périphérie. Les associations sont des centres de régulation médicale qui participent à la permanence des soins ambulatoires en collaboration avec le SAMU (Service d'Aide Médicale d'Urgence). Elles sont regroupées au sein de la Fédération SMF, et participent au système de surveillance syndromique SurSaUD®, que nous présenterons dans la prochaine section.

1.2.4.2. Surveillance syndromique des urgences et des décès

La **surveillance syndromique** permet la surveillance non spécifique d'indicateurs de santé. Avec l'informatisation du système de soins, il est possible de surpasser les délais inhérents à la surveillance basée directement sur les professionnels de santé, et d'élargir le champ de la surveillance sanitaire. Cette surveillance se caractérise par la collecte et le traitement automatiques des données en un temps proche du réel [46]. Les données utilisées sont souvent collectées en routine dans un autre but que celui de la surveillance. Les sources sont variées et se concentrent aussi bien dans le domaine de la santé (vente de médicaments [47], services hospitaliers [48, 49], bases médico-administratives [50, 51], etc.), qu'en dehors (données issues d'Internet, des moteurs de recherche, réseaux sociaux, etc.). En 2011, le panel du projet européen Triple-S définit la surveillance syndromique dans le journal *The Lancet* comme : la

collecte, l'analyse, l'interprétation et la diffusion en temps quasi-réel de données relatives à la santé, pour permettre l'identification précoce de l'impact d'événements sur la santé humaine ou vétérinaire et qui nécessiteraient une action de santé publique. Elle n'est pas fondée sur un diagnostic confirmé, mais sur des indicateurs de santé non spécifiques comme des signes cliniques ou symptômes, eux-mêmes constituant un diagnostic provisoire (ou un « syndrome ») [52]. La surveillance syndromique s'appuie alors sur la détection de motifs aberrants dans les données populationnelles [53]. Pour qu'une source de données soit utile à la surveillance syndromique, sa couverture géographique, démographique et temporelle doit être suffisante pour permettre la détection de signaux [46]. L'avantage majeur de ces sources réside dans leur nature propre. Celles-ci sont presque toujours électroniques, ce qui permet un regroupement par syndromes en temps quasi-réel [46]. Dans cette section, nous présenterons l'utilisation de la surveillance syndromique à partir de sources de données liées à la santé, à travers son application dans la surveillance de la grippe en France. En effet, ce type de surveillance non spécifique est désormais bien intégré au dispositif français de surveillance traditionnel de la grippe. La surveillance syndromique utilisée au-delà des données de santé sera présentée dans le Chapitre 2.

Depuis 2004, le système de **surveillance syndromique SurSaUD®** permet la collecte automatisée de données issues du système de soins. Il complète ainsi en routine la surveillance basée sur la participation directe des professionnels de santé. Cette surveillance non spécifique de la grippe s'effectue à partir de trois sources de données : les urgences en soins primaires à partir de SOS Médecins ; les données d'urgences hospitalières à partir du réseau OSCOUR® ; et les données de mortalité issues de l'Insee (Institut national de la statistique et des études économiques) et du CépiDC (Centre d'épidémiologie sur les causes médicales de décès) de l'Institut national de la santé et de la recherche médicale (Inserm) [33].

Les données de syndrome grippal collectées par SOS Médecins sont transmises quotidiennement à SpF, selon la définition de cas suivante : une fièvre $\geq 38,5^{\circ}\text{C}$ d'apparition brutale, accompagnée de myalgies et de signes respiratoires [54]. Grâce à ces données, SpF estime la proportion hebdomadaire de patients avec un diagnostic de syndrome grippal parmi l'ensemble des actes codés par SOS Médecins aux niveaux national et régional [55].

Depuis 2004, le réseau OSCOUR® (Organisation de la Surveillance COordonnée des URgences) repose sur un réseau de services d'urgences hospitaliers volontaires. Il permet la surveillance du nombre de passages et d'hospitalisations après passage pour différents événements dans les établissements participants, grâce à un recueil de données passif basé sur des regroupements syndromiques. Ceux-ci reposent sur des codes diagnostics de la Classification internationale des maladies – 10ème révision (CIM10). Les indicateurs syndromiques suivis varient au cours de l'année et selon certaines conditions météorologiques. Les données sont ensuite envoyées de façon quotidienne ou hebdomadaire à SpF. Dans le cadre de la surveillance de la grippe, les données de diagnostic codant pour une grippe ou un syndrome grippal sont extraites quotidiennement du dossier médical informatisé, avec les caractéristiques des patients (âge, sexe, lieu de résidence, motif du recours, gravité, orientation de sortie, etc.). L'analyse de ces

données (par classes d'âge et niveau géographique) permet de suivre la dynamique de l'épidémie par l'observation des variations hebdomadaires du nombre de patients consultant aux urgences ou hospitalisés avec un diagnostic de grippe (confirmé ou non). Le suivi de la proportion d'hospitalisations parmi les passages pour grippe constitue un indicateur de gravité de l'épidémie.

En plus de cette surveillance non spécifique à partir des services hospitaliers d'urgence, le dispositif de la grippe français possède un autre indicateur de gravité de l'épidémie de grippe. Depuis 2009, les cas graves de grippe (confirmés ou non) admis en service de réanimation des centres hospitaliers doivent être signalés aux Cellules d'intervention en région (Cire) de SpF, sous forme d'une fiche standardisée. Celle-ci comprend des informations démographiques, sur les facteurs de risque, le statut vaccinal, le résultat virologique, et les éléments de gravité retrouvés chez le patient. Cette surveillance épidémiologique permet d'estimer la gravité de l'épidémie en temps réel et d'identifier les facteurs de risque de grippe grave, afin d'anticiper un éventuel engorgement des structures. Elle est couplée à une surveillance virologique réalisée via les 44 laboratoires du Réseau national des laboratoires hospitaliers (RENAL).

Enfin, la surveillance de la mortalité liée à la grippe s'effectue à partir des certificats de décès qui contiennent deux volets : les données administratives et la cause médicale du décès. Les données administratives permettent une surveillance en temps quasi-réel du nombre de décès total sur le territoire. Ces données s'inscrivent dans le système de surveillance SurSaUD[®], et SpF détermine la part de cette mortalité attribuable à la grippe grâce aux données transmises par l'Insee [31]. Les causes médicales du décès permettent de suivre le nombre de décès directement liés à un événement donné comme la grippe, mais un délai d'obtention de l'information non négligeable est nécessaire. Ce volet est géré par le CépiDC.

1.2.5. Analyse des données de morbidité en soins primaires, l'exemple du réseau Sentinelles

Dans le cadre de la surveillance des maladies infectieuses réalisée en soins primaires, l'indicateur de santé principalement utilisé pour suivre la morbidité dans la population est l'incidence - nombre de nouveaux cas pour un événement de santé donné sur une période définie. Le réseau Sentinelles s'appuie sur la participation de médecins généralistes libéraux pour estimer les incidences de différents indicateurs de santé. Les médecins généralistes Sentinelles (MGS) transmettent régulièrement le nombre de patients vus en consultation, au rythme qui leur convient le mieux (quotidien ou hebdomadaire le plus souvent). Pour que les données collectées ne soient pas trop anciennes et exacerber le biais de mémoire, la période de surveillance peut s'effectuer sur une période rétrospective de 12 jours maximum. La méthode d'estimation est la même pour tous les indicateurs suivis.

Pour estimer le taux d'incidence des indicateurs de santé vus en consultation de médecine générale, la prise en compte de la participation hebdomadaire des MGS et le calcul du nombre

de cas imputés à chaque semaine de surveillance sont d'abord nécessaires. A partir de ces deux valeurs, le nombre hebdomadaire moyen de cas vus par MGS pour une zone géographique est calculé de la façon suivante :

$$\text{Nombre cas moyen (zone, semaine)} = \frac{\sum_{\text{médecins}} \text{nombre de cas (semaine)}}{\sum_{\text{médecins}} \text{participation (semaine)}}$$

Où :

- Le nombre de cas représente le nombre de cas déclarés par un médecin d'une zone donnée sur la semaine
- La participation représente la proportion de jours couverts par la ou les déclaration(s) faite(s) par un médecin pour une semaine donnée (soit sur 7 jours)

L'incidence des syndromes grippaux est alors estimée en extrapolant ce nombre de cas moyen à l'ensemble des médecins généralistes de la même zone géographique. Pour ce faire, deux hypothèses sont avancées :

- Les MGS constituent un échantillon aléatoire représentatif de l'ensemble des médecins généralistes de France, et
- Les MGS déclarent un nombre de cas représentatif de leur activité hebdomadaire.

Pour tenir compte des disparités régionales de densité des MGS, l'incidence est d'abord estimée au niveau régional. Les incidences régionales sont ensuite additionnées pour obtenir l'estimation nationale. Enfin, le taux d'incidence (pour 100 000 habitants) est obtenu en divisant les incidences obtenues pour une zone géographique par la taille de la population annuelle de cette même zone géographique (données de l'Insee) [14, 56].

$$\text{Incidence (France)} = \sum_{\text{régions}} \text{Incidence (région, semaine)}$$

Toutes les données déclarées par les médecins Sentinelles sont validées de façon hebdomadaire par des épidémiologistes. Ainsi, la restitution de la situation des syndromes grippaux observés en *semaine s* par le réseau Sentinelles se fait en *semaine s+1* dans le bulletin de surveillance hebdomadaire.

1.2.6. Avantages et limites de la surveillance traditionnelle de la grippe

Depuis les années 1950, la surveillance épidémiologique basée sur le recueil de données auprès de professionnels de santé a su démontrer son efficacité dans le suivi et la gestion d'épidémies liées à des pathologies infectieuses, mais aussi dans d'autres champs de la santé publique : infections liées aux soins, usage des antibiotiques et résistance bactérienne, maladies chroniques ou la santé liée à l'environnement [1]. Cette surveillance s'effectue sous différentes formes (i) déclaration obligatoire de maladies notifiées, (ii) réseaux de médecins volontaires, ou (iii) notification spontanée en cas d'anomalie perçue, par rapport à la pratique habituelle. Les réseaux

de médecins volontaires en soins primaires sont plus largement représentés. En effet, cette surveillance a l'avantage d'être le premier niveau d'accès de la population au système de soins [57]. Elle permet le suivi spécifique de différents indicateurs de santé, validés par des professionnels et dont la fréquence est importante dans la population (contrairement aux indicateurs suivis par déclaration obligatoire). C'est le niveau de base du système de soins (Figure 2), qui permet de surveiller la santé dans la communauté, auprès d'individus recourant à une consultation en soins primaires.

Malgré tout, cette surveillance n'est donc par définition pas complètement représentative, puisqu'elle ne prend pas en compte la population qui ne consulte pas le système de soins. L'estimation du fardeau réel de l'épidémie peut être sous-évaluée. Par ailleurs, la participation des professionnels de santé repose généralement sur un recueil de données manuel avec souvent une double-saisie d'informations [58]. Les données épidémiologiques collectées sont donc peu détaillées, pour permettre aux médecins de participer plus facilement. L'analyse des données s'effectue ensuite de façon rétrospective. Le délai nécessaire entre la collecte et la restitution des données représente d'ailleurs la limite majeure du système, notamment dans la surveillance des maladies infectieuses [59]. D'autre part, la spécificité et le manque de flexibilité de cette surveillance limite la possibilité de suivre précocement l'émergence d'une pathologie donnée [50]. Chaque système possède sa propre définition de cas, rendant les comparaisons complexes d'un système à l'autre. Enfin, la dernière pandémie grippale a mis en évidence que le recours aux soins peut changer de façon imprévisible lors d'une épidémie inhabituelle. L'augmentation brusque du nombre de cas qui avait recours au système de soins a rendu incertaines les extrapolations des statistiques réalisées et les professionnels étaient débordés, limitant ainsi leur participation au dispositif de surveillance [60].

Pour surpasser ces limites, et grâce aux avancées techniques et technologiques liées au développement d'Internet, des systèmes de surveillance non spécifique ont pu se développer, complétant la surveillance réalisée par les professionnels de santé. La surveillance syndromique basée sur l'extraction de données de santé non spécifiques est désormais bien intégrée aux dispositifs de surveillance, surtout dans les pays développés, comme au Royaume-Uni ou en France [61, 62]. Elle comble les principales limites énoncées ci-dessus de la surveillance basée sur la participation directe de professionnels de santé, avec notamment un suivi quasi-exhaustif selon les systèmes, précis, en temps quasi-réel et sans travail supplémentaire pour le clinicien [63]. Cependant, les sources de données utilisées restent rattachées au système de soins (services hospitaliers, décès, etc.) et ne permettent pas de compléter la surveillance en population générale. Nous détaillerons la surveillance syndromique issue des autres sources de données dans le chapitre suivant (Chapitre 2).

Les messages clés du Chapitre 1

- La surveillance de la grippe est une surveillance globale, complexe et multi-niveaux.
- Les définitions cliniques des cas sont hétérogènes au niveau international et au sein d'un même pays, mais les deux indicateurs principaux sont le syndrome grippal et l'infection respiratoire aiguë.
- Le système de surveillance de référence de la grippe en France s'appuie sur la participation volontaire de professionnels de santé.
- La surveillance syndromique à partir de sources de données de santé est bien établie et complète la surveillance réalisée directement par les professionnels.
- Surveillance basée sur les professionnels de santé → spécifique → délais de collecte et d'analyse des données plus longs
- Surveillance syndromique basée sur le système de soins → non spécifique → collecte et analyse en temps quasi-réel des données
- Les surveillances traditionnelle et syndromique basées sur les données de santé ne prennent pas en compte les personnes qui n'ont pas recours aux soins de santé en cas de maladies grippales.

Chapitre 2.

La surveillance et l'épidémiologie digitale appliquée à la grippe

Dans les années 2000, les maladies infectieuses émergentes (SARS, grippe aviaire A(H5N1), MERS-CoV) ont mis en évidence les lacunes des systèmes de surveillance traditionnelle s'appuyant sur les professionnels de santé, et ont amené à améliorer voire à repenser la surveillance sanitaire, en privilégiant notamment le suivi en temps réel [50, 64]. Aujourd'hui, la surveillance syndromique se développe et s'ouvre plus largement à d'autres sources de données issues d'Internet, des nouvelles technologies et des réseaux sociaux. Les méthodes numériques modernes ont rapidement démontré leur intérêt dans la mise en place de programmes de contrôle et de surveillance des maladies [65]. En 2007, *Ekman et al.* parlent de l'*e-epidemiology* comme un réseau épidémiologique d'individus sentinelles issus de la population pour la surveillance de l'apparition de maladies [66]. *Brownstein et al.* ont ensuite décrit le potentiel d'Internet dans la détection numérique des maladies. Internet permet une vision de la santé mondiale fondamentalement différente de celle qui découle de la déclaration des maladies dans l'infrastructure de surveillance traditionnelle en santé publique [67]. Si on ne retrouve pas de définition consensuelle, *Salathé et al.* définissent l'épidémiologie digitale simplement comme l'épidémiologie qui utilise des données digitales générées à d'autres fins que celles de la santé publique, ce qui est corroboré par *Samerski* qui précise que l'épidémiologie digitale s'appuie sur des données autoproduites [68, 69]. *Eysenbach* parle d'*infodemiology* et d'*infoveillance* pour se référer à l'étude ou à la surveillance des données issues d'Internet à des fins de santé publique [70]. Le point central avec la collecte de données issues d'Internet est que la population devient actrice dans le domaine de la surveillance et de l'épidémiologie digitale : de façon active à travers la surveillance participative (nous la détaillerons dans le Chapitre 3) ou passive avec la surveillance syndromique. Le recueil des données n'est plus limité aux professionnels de santé, mais s'effectue directement auprès de la population, par sa participation volontaire (surveillance participative) ou par les traces laissées sur Internet (surveillance syndromique). La mise en place d'un dispositif de surveillance plus global est alors possible, avec l'inclusion de la population générale qui n'aurait pas recours au système de soins (Figure 4).

L'épidémiologie et la surveillance digitales se sont développées dans différents domaines de recherche, et font l'objet d'études et d'évaluations. Néanmoins, ces études restent presque toujours exploratoires. Dans une revue de la littérature, *Park et al.* ont présenté les différents objectifs de ces études d'épidémiologie digitale. Ceux-ci étaient d'explorer (34,9%), d'expliquer

(27,5%), de décrire (22,9%) et de prédire/contrôler (14,7%) des événements de santé en utilisant Internet comme source de données [4].

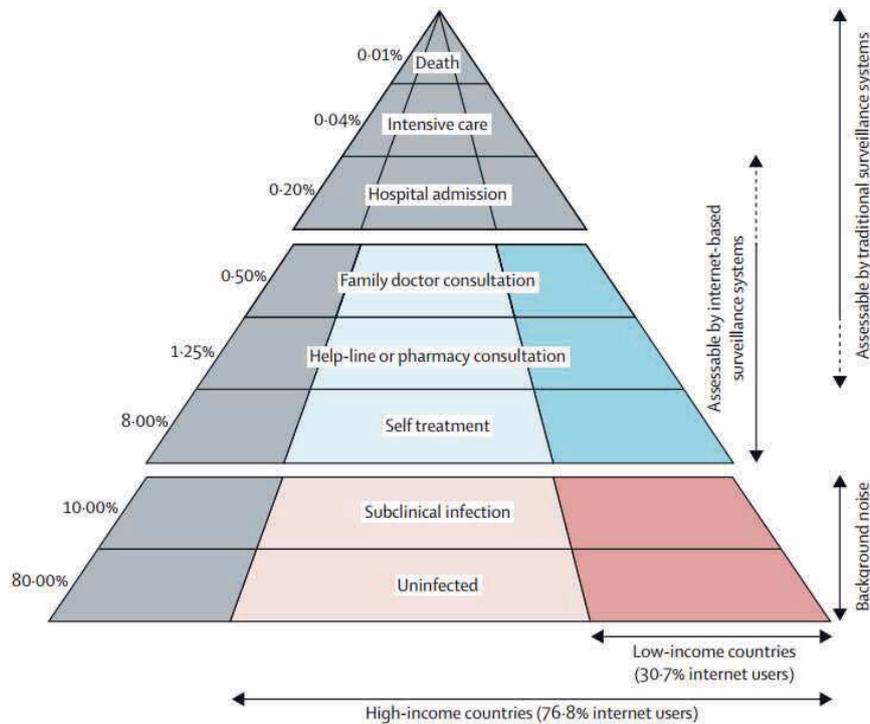


Figure 4. Proportions de la population évaluées par les systèmes de surveillance traditionnels et par Internet pendant une épidémie de grippe dans les pays à revenu élevé et faible. *Source : Milinovich et al. [65] (60)*

2.1. Sources de données et applications

L'utilisation d'Internet comme source d'information pour la mise en place de systèmes de surveillance sanitaire s'est développée dès le début des années 1990. L'intérêt de ces systèmes est la rapidité de la transmission de l'information et de la communication qui en émane. Depuis la création du premier programme de surveillance de maladies infectieuses émergentes en 1994, *Barboza et al.* et *Gajewski et al.* en ont recensé et analysé une dizaine [71, 72]. Les différents systèmes s'appuient sur une ou plusieurs sources de données : médias, pages Internet, salons de discussions, moteurs de recherche, réseaux sociaux, images, données issues des téléphones mobiles, données GPS. En plus des sources de données utilisées, ces systèmes diffèrent par le focus géographique (couverture mondiale ou locale), les langues d'adaptation, le processus de collecte des données (automatique, modération par des professionnels, collecte hybride) et les maladies surveillées. Néanmoins, les performances de détection précoce des événements de santé sont similaires d'un système à l'autre, même si certaines caractéristiques peuvent s'avérer importantes pour une meilleure détection de l'événement surveillé [71]. Par exemple, les systèmes modérés par des professionnels de santé (ProMED-mail, GPHIN, HealthMap) permettent d'assurer une meilleure capacité de détection des événements. Cette dernière

caractéristique est d’ailleurs exploitée dans les systèmes participatifs qui se sont développés à partir des années 2000, pour un suivi un peu plus spécifique. Nous les détaillerons dans le Chapitre 3.

Une revue d’études épidémiologiques utilisant des sources de données digitales a montré que le domaine le plus étudié était celui des maladies infectieuses. Sur les 109 études incluses, près de 60% des études concernaient les maladies infectieuses, et la grippe représentait plus d’un tiers des études menées sur les maladies infectieuses [4]. Aujourd’hui, environ 65% des premières informations mondiales relatives aux maladies infectieuses proviennent de sources informelles comme Internet. La majorité des épidémies investiguées par l’OMS sont donc identifiées en premier lieu par ces sources de données [72, 73].

2.1.1. Description des principaux systèmes de surveillance utilisant Internet comme source d’information

Parmi les différents systèmes de surveillance en place, trois sont principalement retrouvés dans la littérature. En 1994, ProMED-mail a été le premier programme de surveillance des maladies émergentes développé, mais GPHIN et HealthMap sont actuellement les plus répandus. Les trois systèmes sont caractérisés dans le Tableau 1.

Tableau 1. Caractéristiques de trois systèmes traitant les sources de données informelles basées sur des événements. Source : Adaptation à partir de Keller et al. [74]

System	Data sources (languages)	Data characterization	Information formatting	Data dissemination		
				Access	User interface	Format
GPHIN	Factiva, Al Bawaba (9 languages)	Automatic and human	Categorization, machine translation, geocoded	Subscription only	Boolean and metadata query system (native)	Email alert
HealthMap	Google news, Moreover, ProMED, WHO, Eurosurveillance (4 languages)	Automatic	Categorization, geocoded, time coded, extra information	Open	Mapping, faceted browsing (native)	RSS feed
ProMED-mail	Subscribers, Medias, Official reports (7 languages)	Human	No formatting	Subscription only (public health workers and general population)	Reports	Email alert, RSS feed

ProMED-mail est un système de déclaration et de diffusion d’alertes sanitaires par mail en rapport avec la santé humaine, animale et végétale [67, 75, 76]. L’objectif est d’avoir un système rapide de transmission avec une large couverture de diffusion (scientifiques, médecins, épidémiologistes, acteurs de santé publique et population générale). Ce système est ouvert à toutes les sources de données en ligne (médias, rapports officiels, observations locales, etc.). Les alertes qui en découlent sont souvent rapportées par les abonnés, et modérées avant publication sur le réseau. Elles sont alors diffusées par mail à l’ensemble des abonnés (~70 000 dans 185 pays), et affichées sur le site Internet. ProMED-mail a démontré son utilité pour la surveillance des maladies émergentes, notamment avec l’épidémie de SRAS (Syndrome Respiratoire Aigüe Sévère) de 2003 [67] et du MERS-CoV (coronavirus du syndrome respiratoire du Moyen-Orient) de 2012 [77], en informant précocement les responsables de santé publique. Le système avait identifié des rapports textuels informels (médias, salons de discussion) sur le Web. La Figure 5 montre la localisation des abonnés et les différentes maladies recensées par ProMED-mail en fonction des régions du monde [76].



Figure 5. Localisation des modérateurs et abonnés (à gauche) et maladies recensées selon les régions du monde entre janvier 2016 et décembre 2016 (à droite) par le système ProMED-mail. *Source : Carrion et al. [76]*

Dans la lignée de ProMED-mail, le GPHIN (Global Public Health Intelligence Network), initié par l’agence de santé publique canadienne et l’OMS, a été développé en 1997 [78, 79]. Ce système est basé sur un logiciel qui récupère des informations informelles sur Internet, permettant de repousser les limites de la surveillance sanitaire institutionnelle au-delà des frontières nationales et d’apporter une réponse globale. Le GPHIN a démontré son efficacité dans la détection précoce de menaces émergentes, comme en 2003 avec le SARS, et cela bien avant les déclarations officielles [80]. Il a alors permis de renforcer la surveillance sanitaire et les réponses de l’OMS sur la scène internationale, en permettant à l’OMS d’appuyer ses efforts menés auprès des Etats Membres à déclarer au plus tôt les épidémies dans leur pays. La différence entre ProMED-mail et GPHIN réside dans le fait que le dernier ne se base que sur les sources médiatiques internationales filtrées par le logiciel et non sur le rapport d’événements par les abonnés. D’autre part, le GPHIN n’est pas public, seuls les professionnels de santé publique y ont accès.

En 2006 le projet HealthMap a été lancé. Il s'agit d'un système gratuit, multilingue, qui permet l'extraction de contenus textuels à partir de sites Internet pour la surveillance de maladies infectieuses. HealthMap permet l'identification d'événements de santé (comme les épidémies), d'agents pathogènes et de leur localisation géographique [81-83]. Les informations collectées sont représentées dans un système d'information géographique sous la forme d'une carte mondiale interactive. HealthMap utilise par exemple comme source de données des moteurs de recherche comme Baidu ou Google, les données issues de Eurosurveillance, ProMED-mail ou de l'OMS. Ce système permet d'avoir une vue synthétique de l'ensemble des phénomènes au niveau mondial et en temps réel, en dehors des systèmes de surveillance traditionnelle. HealthMap a permis l'alerte précoce et le suivi de pathologies infectieuses avec succès, comme lors de la pandémie grippale A(H1N1) de 2009, ou plus récemment lors des épidémies de MERS-CoV ou Zika [84-86].

La plupart des principaux systèmes de surveillance, qui usent d'Internet au sens large comme source d'information, ont été développés pour la détection précoce des maladies infectieuses (comme les maladies émergentes). La surveillance d'indicateurs de santé en routine n'est pas leur principal objectif. Ainsi, seul HealthMap mène des projets sur la surveillance de la grippe, comme HealthMap Flu Trends [87] ou HealthMap Vaccine Finder [88]. HealthMap Flu Trends est une approche de modélisation par apprentissage automatique pour prédire l'activité grippale grâce à différentes sources de données : Google, Twitter, les dossiers médicaux électroniques, et les données issues du système de surveillance participatif Flu Near You (voir Chapitre 3). Elle fournit des estimations en temps réel et permet la prévision des incidences des syndromes grippaux aux niveaux national et régional aux USA, jusqu'à deux semaines avant les rapports du CDC. HealthMap Vaccine Finder permet de mettre en relation la population américaine et les fournisseurs de vaccins, ainsi que de communiquer sur les bienfaits de la vaccination. La disponibilité des vaccins est présentée sur une carte interactive générant les stocks autour des utilisateurs. Une personnalisation des vaccinations recommandées peut être réalisée pour les usagers.

Les systèmes utilisant Internet comme source principale d'information tendent à évoluer pour affiner la surveillance réalisée. Leur amélioration passe par l'inclusion de nouvelles sources de données, comme l'utilisation des moteurs de recherche et des médias sociaux dans le projet HealthMap Flu Trends, par le perfectionnement des techniques d'analyse ou de visualisation des données. Le suivi de la grippe et des syndromes grippaux a souvent servi de cas d'étude pour l'utilisation des moteurs de recherche et des médias sociaux comme sources supplémentaires de données de surveillance [89].

2.1.2. Les moteurs de recherche comme source d'information

A la fin des années 2000, des études ont proposé l'utilisation des requêtes Internet issues des moteurs de recherche pour la surveillance épidémiologique. En effet, un nombre important de

personnes utilisent Internet pour obtenir des informations sur leur santé avant de consulter un professionnel, même si d'importantes disparités persistent concernant l'utilisation d'Internet entre les pays aux revenus faibles et ceux aux revenus élevés (Figure 6). Aux USA seulement, plus de 8 millions de requêtes Internet relatives à la santé sont générées chaque jour [65]. Les recherches effectuées par les utilisateurs concernant des symptômes grippaux et enregistrées via différents moteurs de recherche ont été comparées à des données issues des systèmes traditionnels. Les premières études ont été menées aux USA en comparant les données liées à la grippe issues du moteur de recherche Yahoo à celles du CDC (cas de grippe confirmés et mortalité par pneumopathie et grippe) [90]. L'objectif était de prédire le pourcentage de cultures positives pour la grippe et les décès attribuables aux pneumopathies et à la grippe aux USA grâce aux requêtes Yahoo. Une prédiction de l'augmentation des cas positifs et de la mortalité liée à la grippe a été réalisée une à cinq semaines avant leur détection par le système de surveillance traditionnel [90]. *Ginsberg et al.* ont ensuite utilisé les requêtes Internet de Google pour construire un modèle prédictif de la grippe aux USA, Google Flu Trends (GFT) [91]. Le système s'est révélé efficace dans le suivi des symptômes grippaux en complément des données du CDC et a été développé dans près de 29 pays. L'augmentation des requêtes liées à la grippe avait lieu une à deux semaines avant la détection faite par le système traditionnel et à un niveau plus local. Ce modèle a ensuite été calibré sur les données grippe du CDC afin de réduire le bruit inhérent à ce type de données de masse. En effet, la grippe n'est que très indirectement mesurée car les recherches effectuées ne signifient pas forcément que la personne est grippée. Malgré cette calibration, GFT n'a pas réussi à détecter la première vague pandémique de 2009 liée à la temporalité et aux caractéristiques inhabituelles de l'épidémie [7, 92, 93]. Lors de la saison grippale 2012/2013, GFT a cette fois surestimé la sévérité de la maladie par rapport aux observations faites par les systèmes traditionnels [7, 93]. Google a alors abandonné le suivi de ses requêtes Internet, en laissant ses algorithmes à la disposition du CDC et d'autres institutions [89, 94]. En parallèle de ses études menées aux USA avec le moteur de recherche Google, des études similaires ont été menées en Chine avec le moteur de recherche Baidu, plus utilisé. Une très bonne corrélation des données liées à la grippe était observée entre le moteur de recherche et les sources institutionnelles, avec des estimations plus précoces pour Baidu [95]. L'encyclopédie universelle Wikipédia a également été utilisée à des fins de veille sanitaire, en rendant disponible le nombre d'articles vus par heure [96, 97]. Le taux de consultation quotidien de certains articles Wikipedia liés à la grippe a été comparé aux données issues du CDC. Le système était efficace pour suivre les épidémies de grippe, et moins sensible aux caractères inhabituels retrouvés lors de la pandémie [98].

Au-delà de la grippe, les moteurs de recherche ont été utilisés pour le suivi d'autres maladies infectieuses comme les gastroentérites [99], la dengue [100, 101], le paludisme [102] ou encore la maladie de Lyme [103, 104], ou pour le suivi d'indicateurs de santé non-infectieux comme le risque suicidaire [105]. Google Trends a permis aussi d'évaluer l'intérêt de la population sur des sujets de santé comme le cancer [106] ou d'évaluer l'impact de campagnes de communication sanitaires comme l'évaluation des connaissances de la population sur les soins palliatifs [107].

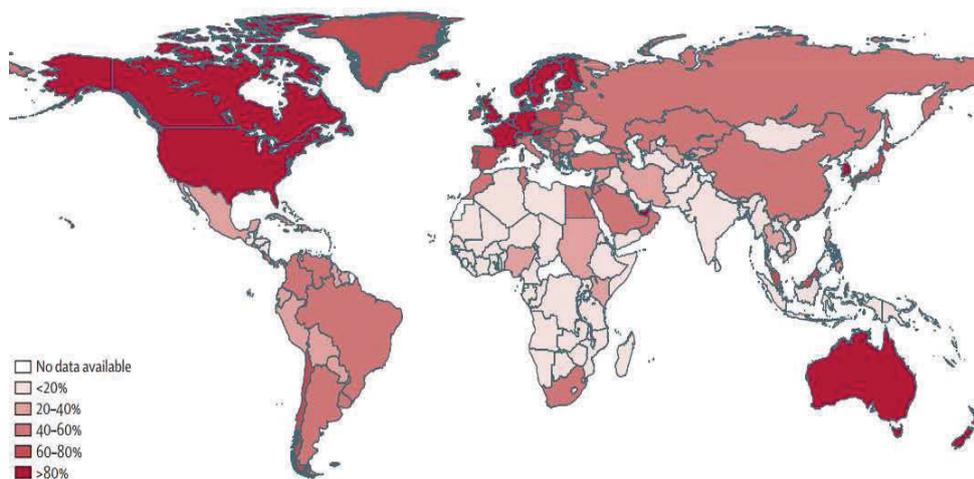


Figure 6. Utilisation d'Internet par pays (en % de population). *Source : Milinovich et al. [65]*

2.1.3. Les réseaux sociaux comme source d'information

En parallèle de l'utilisation des moteurs de recherche à des fins de veille sanitaire, les réseaux sociaux représentent également une source d'information et de diffusion de plus en plus reconnue en santé publique. Ils sont un moyen de communication dynamique et interactif, avec un fort taux de pénétration en population générale, dans les pays à haut et moyen revenus [108]. Les premières études ont analysé le contenu des blogs, avec l'extraction de certains mots-clés et la comparaison avec des indicateurs de santé issus des systèmes traditionnels [109]. De tous les médias sociaux, les microblogs, et particulièrement le réseau Twitter a connu un essor important à partir des années 2010 dans la recherche en santé. Twitter permet de publier de courts messages textuels, ou de relayer des messages d'autres utilisateurs en temps réel. Une revue de la littérature a montré que le contenu des tweets était analysé dans le champ de la surveillance sanitaire, et à moindre échelle pour exploiter ses fonctions interactives pour le recrutement ou les interventions [110]. Il a d'abord été utilisé pour caractériser la pandémie grippale A(H1N1) de 2009, afin d'étudier l'intérêt de ce type de données dans la surveillance de la grippe. Des études ont été menées pour comparer la corrélation entre les données issues de Twitter et celles du CDC aux USA [111, 112], et pour mesurer l'intérêt et les préoccupations des utilisateurs face à la grippe pandémique [113]. Comme pour les moteurs de recherche, cette source de données permet d'anticiper l'activité grippale d'une à deux semaines. Des études complémentaires ont été réalisées pour le suivi des gripes saisonnières. Là encore, de bonnes corrélations ont été observées avec les systèmes traditionnels. Par ailleurs, Twitter permet le suivi à une échelle plus fine (région, ville) et performe mieux que GFT pour le suivi de la grippe localement [114-116]. Cependant, des études plus récentes comparant Twitter et Google ont mis en évidence certaines lacunes du réseau social. Twitter s'avèrerait peu sensible avec une faible valeur prédictive positive (VPP), tandis que Google est très sensible avec une VPP modérée [97]. De plus, Twitter a une plus faible imprégnation dans la société par rapport à Google [117].

2.2. Avantages et limites

L'utilisation et l'étude des différents médias ayant recours à Internet ont mis en évidence les qualités de ces nouveaux systèmes par rapport aux systèmes existants. La qualité majeure largement retrouvée dans les diverses études concerne la détection précoce et l'anticipation des événements de santé par rapport aux systèmes s'appuyant sur les professionnels de santé [67, 83, 118-121]. La sensibilité des systèmes permet la détection précoce d'épidémies et la connaissance en amont de la localisation géographique de l'événement de santé, facilitant les réponses des acteurs de santé publique [67, 83, 122]. L'anticipation d'événements, comme la connaissance du pic épidémique d'une épidémie de grippe, permet de planifier la gestion logistique du système de soins pour éviter son engorgement en cas d'afflux, par la communication faite par les institutions sanitaires auprès des professionnels et du grand public à une échelle nationale ou plus locale [83]. Par ailleurs, ces systèmes basés sur la surveillance syndromique permettent un recueil de données plus large et systématique, directement à partir de données générées par la population générale (Figure 4). Cela permet de compléter la surveillance des maladies en collectant des informations auprès des personnes qui ne consultent pas le système de soins, de personnes vulnérables qui ne seraient pas prises en compte dans les systèmes traditionnels, et de mener des actions de santé publiques ciblées pour améliorer l'état de santé de la population [120, 122-124]. Enfin, l'exploitation de ces données peut se faire à faible coût sur de grands échantillons [66, 83]. Le suivi de l'évolution des événements de santé est d'ailleurs facilité par l'automatisation et l'uniformité de la collecte des données en temps réel à différents niveaux (national ou international), aussi bien pour des événements de santé infectieux que non-infectieux [46, 118]. Les données n'étant pas spécifiquement conçues pour le suivi épidémiologique, il n'y a pas de biais de réponse de la part des personnes pour être socialement convenables [66, 120].

Malgré les qualités observés, différentes limites ont été mises en évidence concernant l'utilisation des données digitales extraites d'Internet comme sources de données épidémiologiques. La limite majeure observée concerne la calibration des données sur les systèmes de surveillance traditionnelle [7, 91, 92]. Cette calibration est donc faite sur des données déjà connues. La valeur ajoutée de ce type de sources de données est donc faible si on prend l'exemple du suivi des maladies infectieuses. Les sources de données basées sur la surveillance participative, que nous décrivons dans le Chapitre 3, représentent donc une meilleure alternative pour étudier le phénomène directement à partir de la population. Par ailleurs, les données issues d'Internet sont sensibles aux paramètres extérieurs comme la couverture médiatique [67]. En cas de couverture importante d'un phénomène de santé, le nombre de requêtes Internet augmente sur ce sujet. Les personnes se renseignent ou veulent connaître la situation via Internet, alors qu'elles ne sont pas forcément directement impactées par la maladie, comme ce fut le cas lors de la pandémie de 2009 [125]. L'extrapolation des données et le lien de causalité avec l'événement de santé donné ne sont donc pas aisés. D'autre part, la nature « incontrôlée » des données enregistrées ainsi que le manque de validation (i.e. collecte automatique) représentent un enjeu important à la

précision et à la bonne estimation de l'événement de santé par le système [119, 124]. Enfin, l'évolution rapide de ces sources de données massives rend difficile leur analyse et gestion. D'ailleurs, si le suivi en temps réel est un atout majeur, les études réalisées ont évalué les signaux obtenus par ces sources de données de façon rétrospective [120, 126].

Les messages clés du Chapitre 2

- Les données issues d'Internet (moteurs de recherche, réseaux sociaux, blogs) ont démontré leur intérêt dans le champ de la santé publique, et particulièrement dans le suivi d'événements inhabituels ou d'épidémies.
- D'un côté, la surveillance digitale permet la détection et l'alerte précoce en cas de l'émergence de nouvelles maladies. De l'autre, la rapidité de la collecte et de l'analyse des données permet la surveillance continue d'épidémies en cours, et d'anticiper les signaux observés par les systèmes traditionnels.
- Ces données permettent donc de compléter les données traditionnelles, en captant la population qui n'a pas recours au système de soins. Cependant, les données issues d'Internet sont sensibles aux événements extérieurs et souvent calibrées sur des données existantes, limitant leur intérêt par rapport à la surveillance participative.
- Les études menées à partir de ces sources de données de surveillance syndromique restent exploratoires, sans réflexion autour de leur intégration aux dispositifs de surveillance existants.

Chapitre 3.

Sciences participatives et surveillance des maladies infectieuses

3.1. Principes de la surveillance participative et applications

La surveillance syndromique utilisant les sources de données issues de la population générale via Internet s'effectue soit par une collecte passive, comme nous l'avons vu dans le Chapitre 2, soit par la collecte active de données de santé comme nous allons le voir dans ce Chapitre. Depuis le début des années 2000 et surtout après la pandémie grippale de 2009, des nouveaux systèmes participatifs ont été développés toujours dans l'idée de compléter la surveillance réalisée par les systèmes traditionnels. La participation est presque toujours basée sur le volontariat, et le recueil de données est plus spécifique avec des données structurées. Différents termes sont utilisés pour décrire ces systèmes participatifs sans qu'il n'y ait de consensus quant à leur nomination. Les systèmes font appel à trois concepts contemporains : les sciences participatives, le *crowdsourcing* et la *eHealth* [127]. Les sciences participatives sont la conduite d'activités scientifiques par des personnes qui n'ont pas de formations spécifiques dans le domaine étudié [127]. Même si le concept n'est lui pas nouveau, le terme de *crowdsourcing* (ou production participative) a été conçu en 2006 par Jeff Howe, rédacteur au Wired magazine, pour désigner le recrutement d'une main-d'œuvre mondiale (services, idées, contenu) via Internet sans avoir recours à de la sous-traitance [128]. La *eHealth*, quant à elle, considère le déploiement des outils Web 2.0 dans le domaine de la santé et la participation active des individus considérés comme des partenaires des acteurs de soins [127]. La différence entre sciences participatives et *crowdsourcing* réside dans le fait que ce dernier n'implique pas forcément la participation de non-spécialistes. Il peut utiliser la technologie *eHealth*, sans que cela ne soit une exigence [129]. D'autres définitions du *crowdsourcing* ont été décrites, bien que similaires. *Boulos et al.* parlent du *crowdsourcing* comme d'un moyen de capitaliser sur le pouvoir des masses et de s'appuyer sur la participation citoyenne pour atteindre un objectif [130]. *Ponsonby et al.* le définissent comme un travail de groupe caractérisé par l'agrégation des décisions individuelles, prises en compte de manière indépendante [131]. Le principe général de la recherche scientifique basée sur les sciences participatives et le *crowdsourcing* est de s'appuyer sur l'intelligence collective qui émerge lorsque des personnes habituellement non impliquées dans un domaine sont mobilisées pour travailler autour d'un objectif commun [132].

Différents domaines d'application ont été étudiés, mais la promotion de la santé, avec notamment la surveillance, les connaissances et l'éducation à la santé, a été la plus exploitée jusqu'à présent [129, 133]. La surveillance participative est une forme de sciences participatives, qui se base sur le principe que la population générale est la première source d'information pour la mise en place d'études épidémiologiques à grande échelle [134, 135]. L'idée est de tirer parti de la connectivité numérique pour inciter le public à fournir activement aux acteurs de santé publique des données qui peuvent être regroupées et analysées à diverses fins [136]. Dans le cadre de la surveillance participative, l'intérêt est de comprendre la dynamique des maladies directement à partir de la population générale, car cela représente la meilleure opportunité de comprendre, contrôler et prédire la diffusion de la maladie [137].

La surveillance participative est surtout appliquée à des événements aigus comme le suivi des maladies infectieuses (émergentes ou non), le suivi de menaces sanitaires comme la sécurité alimentaire ou les dangers environnementaux, et les événements liés à la santé animale [136]. En dehors de l'aspect infectieux, d'autres champs de promotion de la santé publique ont été étudiés par le biais de la surveillance participative, afin de mieux connaître la situation de l'indicateur en question, de suivre son évolution et de mettre en place des actions de santé publique ciblées. Cela concerne surtout les addictions (alcool, tabac, nutrition), les rapports sexuels à risque ou la santé mentale [129, 133]. Le Tableau 2 détaille les principaux systèmes de surveillance participatifs retrouvés dans le monde pour le suivi des maladies infectieuses. La surveillance participative de la grippe est plus fréquemment étudiée. En 2003, les Pays-Bas ont été précurseurs à la mise en place de ce type de surveillance. Des systèmes similaires se sont rapidement développés dans d'autres pays d'Europe, mais aussi en Australie, USA, Porto Rico, Brésil, Thaïlande, Canada, Tanzanie.

Au-delà de la grippe, les maladies infectieuses concernées par la surveillance participative sont la dengue (Brésil, Porto Rico, Sri Lanka, USA) ou le chikungunya, et des maladies infectieuses émergentes comme Zika (Tableau 2). Cela permet d'avoir une vision locale de la diffusion de ces maladies en population générale.

Tableau 2. Principaux systèmes de surveillance participatifs à travers le monde, pour le suivi de maladies infectieuses. Source : Adaptation à partir de Leal-Neto et al. [134] et de Smolinski et al. [136]

Name	Country, base year	Health sector	Disease focus	Registered users	Reporting frequency	Modality
InfluenzaNet [138-140]	NL, 2003-2017 [141]					
	BE, 2003-2017 [141]					
	PT, 2005 [142]					
	IT, 2008 [143]		ILI ¹		Weekly	
	UK, 2009 [144]				(incentive	Website,
	SE, 2011 [145, 146]	Human	SE : global health	350-50,000	through	email,
	FR, 2012 [147, 148]		watch since 2015		weekly	mobile app
	ES, 2012		(Hälsorapport)		newsletters)	
	IE, 2013					
	DK, 2013 [149]					
CH, 2016						

Flutracking [150]	AU, 2006	Human	ILI ¹	400-30,900	Weekly	Website, email
Flu Near You [151, 152]	USA, 2011	Human	ILI ¹	>50,000	Weekly	Website, email, mobile app
GrippeWeb [153]	DE, 2011	Human	ILI ¹ , ARI ²	5000	Weekly	Website, email
Dengue na Web	BR, 2011	Human	Dengue	-	-	Website
Salud Boricua	PR, 2012	Human	ILI ¹ , dengue, leptospirosis, chikungunya	500-5000	Weekly	Website
Saude na Copa [154] (Healthy Cup)	BR, 2014	Human	ILI ¹ , measles, rubella, cholera, acute diarrhea, dengue	500-5000	Daily	Website, mobile app
DoctorMe	TH, 2015	Human	ILI ¹ (+ personal health care guidance)	15,000-50,000	Daily	Mobile app
Participatory One Health Disease Detection [155]	TH, 2015	Human, Animal, Environment	All syndromes, livestock outbreaks, natural disasters	300	Event-based	Mobile app
FluWatchers	CA, 2015	Human	ILI ¹	<500	Weekly	Website, email
Guardioes da Saude (Guardians of Health)	BR, 2016	Human	Acute respiratory exanthematic syndromes and diarrhea	500-5000	Daily	Website, mobile app
Mo-Buzz [156]	LK, 2016	Human, Environment	Dengue, mosquito-breeding site	<500	Daily	Mobile app
AfyaData [157]	TZ, 2016	Human, animal, environment	All syndromes, livestock outbreaks, wildlife outbreaks	<500	Event-based	Website, mobile app
Kidenga [158]	USA, 2016	Human, environment	Zika, chikungunya, dengue	<500	Weekly	Mobile app

¹Influenza-like-illness

²Acute respiratory infection

Dans la section suivante, nous allons présenter l'exemple de la plateforme participative européenne InfluenzaNet. Nous détaillerons ensuite plus spécifiquement la cohorte française GrippeNet.fr.

3.2. InfluenzaNet ou la surveillance participative des syndromes grippaux en Europe

Le projet d'une surveillance participative en population à un niveau européen est né de l'expérience positive menée aux Pays-Bas (et à la partie flamande de la Belgique) (degrotegiepmeting.nl) et au Portugal (www.gripenet.pt), respectivement en 2003 et 2005. Les deux pays ont mis en place un site Internet permettant le recueil de données épidémiologiques liées à la grippe directement à partir de la population générale. Face au succès reçu, 20 000 participants à l'issue de la première année aux Pays-Bas et environ 5 000 participants sur les deux premières années au Portugal, le projet s'est développé en Italie en 2008 (www.influweb.it). En 2009, les quatre pays ont alors formé le projet EPIWORK (financé par l'Union Européenne) qui visait à mettre en place une plateforme Internet européenne. InfluenzaNet est née de la collaboration de différents acteurs de santé publique européens (agences de santé publique, universités et instituts de recherche, entreprises privées), aux compétences variées (épidémiologistes, médecins de santé publique, modélisateurs). Les données épidémiologiques collectées et les plateformes techniques ont été standardisées pour permettre une surveillance en population à l'échelle européenne, afin de pouvoir comparer facilement le suivi des indicateurs grippaux (syndrome grippal et IRA) entre les pays. Les données collectées portent sur différents symptômes, sur l'accès et l'utilisation du système de soins, l'automédication, la vaccination antigrippale et ses déterminants, et l'absentéisme en cas de grippe. La volonté d'étendre le système à d'autres pays s'est précipité avec la pandémie grippale A(H1N1) de 2009. Le Royaume-Uni a rejoint le projet en 2009 (www.flusurvey.net), la Suède en 2011, la France (www.grippenet.fr) et l'Espagne (www.gripenet.es) en 2012, le Danemark (www.influmeter.dk) et l'Irlande (www.flusurvey.ie) en 2013, et la Suisse en 2016 (www.grippenet.ch). Au total, près de 40 000 personnes participent chaque hiver à la surveillance de la grippe dans les dix pays européens.

D'un point de vue légal et éthique, chaque pays applique ses lois nationales relatives à la vie privée, à la collecte et au traitement des données [138-140]. Depuis mai 2018, le règlement général sur la protection des données (RGPD), mis en place par la Commission Européenne, est appliqué collectivement à l'ensemble des plateformes intégrées à InfluenzaNet. *Geneviève et al.* ont d'ailleurs proposé un cadre pour réglementer et renforcer les approches éthiques des systèmes participatifs (i.e. consentement, protection de la vie privée, etc.), en prenant l'exemple d'InfluenzaNet pour illustrer les défis éthiques liés à ces systèmes de collecte de données [159]. Un système sécurisé et fiable permet d'accroître la confiance des participants et par conséquent d'assurer leur participation.

3.2.1. Collecte et gestion des données

La participation à InfluenzaNet est volontaire et bénévole, seule une adresse e-mail est requise. Le recrutement des participants s'effectue par des moyens de communication propre au pays. Il peut s'agir de relais réalisés par les institutions sanitaires qui participent au projet, d'articles, d'interviews ou de reportages dans les médias, de diffusions via Internet ou les réseaux sociaux, et du bouche-à-oreille réalisé par les participants. Les participants ne s'inscrivent pas à la plateforme InfluenzaNet à proprement parlé, mais à la plateforme du pays européen dans lequel ils vivent, qui est membre du consortium InfluenzaNet.

Le suivi a lieu en général entre octobre/novembre et avril/mai. En début de saison de surveillance, les participants répondent à des questions sociodémographiques, médicales et comportementales. Ce questionnaire d'inclusion peut être actualisé au cours du suivi, notamment avec la mise à jour du statut vaccinal antigrippal (cf. Annexe 1). Une personne peut participer pour un autre membre de son foyer ou extérieur à celui-ci. Ceci afin de pouvoir collecter des données auprès de personnes non initiées à l'usage d'Internet (les plus jeunes et les plus âgées). Après quoi, les participants sont invités chaque semaine à signaler la présence ou l'absence de symptômes depuis leur dernière connexion, à partir d'une liste de symptômes facilement identifiables : fièvre, frissons, rhinorrhée ou congestion nasale, éternuements, mal de gorge, toux, dyspnée, maux de tête, myalgie, douleur thoracique, fatigue ou épuisement (malaise), perte d'appétit, expectorations colorées, yeux larmoyants ou rouges, nausées, vomissements, diarrhées, maux d'estomac, ou d'autres symptômes (cf. Annexe 2). La (Figure 7) détaille le processus de collecte des données. C'est après la pandémie grippale de 2009 que la liste s'est élargie pour inclure des symptômes gastro-intestinaux. L'objectif était d'obtenir une liste complète qui permet d'étudier aussi d'autres maladies potentiellement liées à la grippe. Si un symptôme est déclaré, des questions supplémentaires sont posées pour mieux le caractériser. La combinaison de plusieurs symptômes déclarés (comme la fièvre, les frissons, la toux, etc.) et les détails de leur survenue (degré de température de la fièvre, soudaineté des symptômes) permettent d'évaluer le nombre de cas observés avec un épisode grippal (syndrome grippal ou IRA), mais aussi de caractériser le comportement des participants en cas de maladie (recours aux soins, prise de médicaments avec entre autres les antiviraux et les antibiotiques, absentéisme). Différents symptômes ont été retenus afin de pouvoir simuler toutes les définitions de cas de syndrome grippal et IRA retrouvées dans chacun des pays participants. La définition de référence au sein d'InfluenzaNet est la définition de syndrome grippal de l'ECDC, à savoir : soudaineté des symptômes ET au moins un signe systémique parmi la fièvre ou les frissons, le malaise, les maux de tête, ou les myalgies ET au moins un signe respiratoire parmi la toux, les maux de gorge ou la dyspnée [160]. La reconstruction d'autres définitions de cas est bien sûr possible grâce à cette surveillance basée sur la collecte de simples symptômes. La définition de l'OMS d'un syndrome grippal et d'autres définitions propres aux pays ont été explorées [138, 148, 149].

Les données sont ensuite restituées au niveau européen sur le site Internet d'InfluenzaNet et spécifiquement sur les sites des différents pays adhérents. Un projet plus large FluOutlook a été mis en place dernièrement. Ce projet s'appuie sur l'analyse des données issues d'InfluenzaNet par des modèles développés dans la plateforme Gleam, pour réaliser des simulations et des prévisions. L'objectif est d'étudier la dynamique de transmission des épisodes grippaux dans les pays, d'en estimer les paramètres épidémiologiques clés et de prévoir la tendance à long terme de l'épidémie [87].

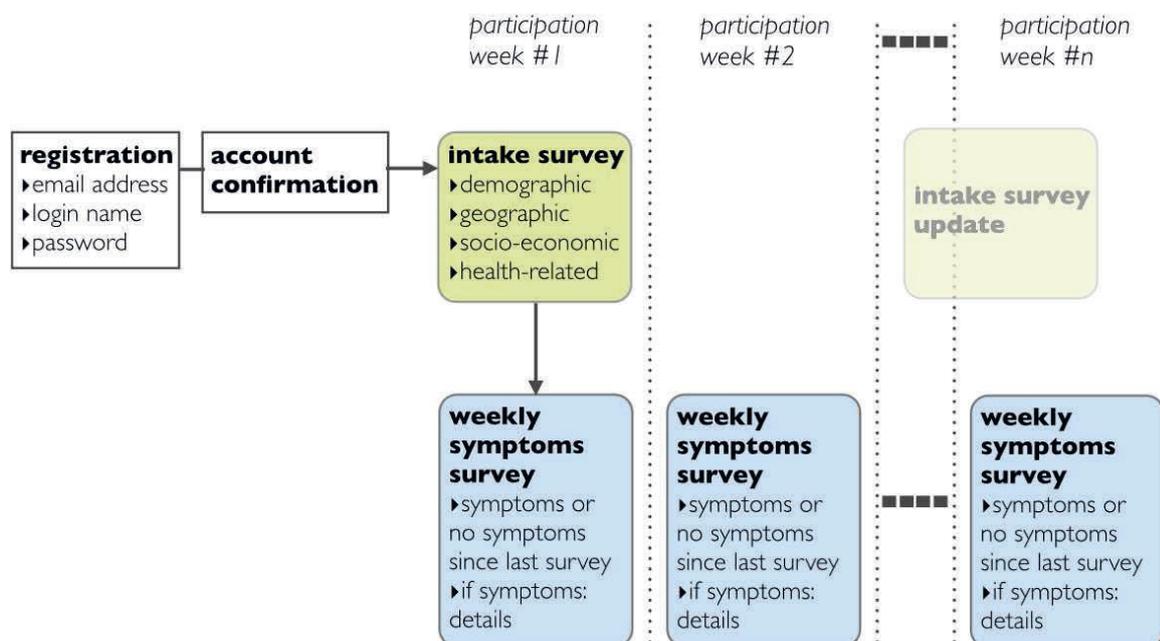


Figure 7. Représentation de la collecte des données entre octobre/novembre et avril/mai au sein de la plateforme InfluenzaNet. *Source : Cantarelli et al [161].*

3.2.2. Représentativité des données collectées

Comme attendu, des différences statistiquement significatives sont retrouvées lorsque l'on compare la population inscrite à InfluenzaNet à la population générale des pays. En effet, par principe, un biais de sélection était attendu puisque les cohortes participatives sont ouvertes au volontariat et non échantillonnées à partir de la population générale. Un travail européen pour mieux comprendre les caractéristiques des participants a été mené à partir des données de 2011/2012, dans les sept premiers pays adhérents (Pays-Bas, Belgique, Portugal, Italie, Royaume-Uni, Suède et France) [161]. De façon globale, les participants se sentent concernés par la thématique de la grippe et par leur santé de façon générale (moins de fumeurs, plus de personnes vaccinées notamment chez les 65 ans et plus pour qui la vaccination est recommandée). Ils sont d'ailleurs en meilleur état de santé que le reste de la population. En détails :

- Les femmes étaient surreprésentées dans la majorité des pays de la cohorte (Pays-Bas, Royaume-Uni, Suède et France), mais on retrouvait plus d'hommes en Italie et en

Belgique. Ceci pourrait s'expliquer par le fait que les femmes sont plus actives que les hommes lorsqu'il s'agit de rechercher des informations relatives à leur santé [162].

- Toutes les classes d'âge étaient représentées, mais les adultes de 40 à 69 ans étaient surreprésentés, alors que les plus jeunes (0 – 30 ans) et les plus âgés (≥ 70 ans) étaient sous-représentés. Ceci est souvent observé avec la collecte de données sur Internet [120, 124]. Néanmoins, une augmentation de la participation des personnes plus âgées a été observée dans la cohorte avec le temps, sûrement liée à une utilisation accrue d'Internet [163].

Des études de représentativité similaires ont été menées spécifiquement par certains pays adhérents, pour mettre à jour les données ou parmi les nouveaux pays intégrés non couverts par l'étude européenne [144, 146, 147, 149].

3.2.3. Incidence des syndromes grippaux estimée par les systèmes participatifs en Europe

L'estimation des incidences à travers InfluenzaNet s'effectue de façon hebdomadaire, par le rapport du nombre de cas de syndrome grippal par semaine sur le nombre total de participants pour cette même semaine [138, 144, 146, 148, 149, 163]. Certains travaux ont rapporté ce nombre à la population générale en pondérant les données par âge et/ou sexe [146, 148]. L'estimation des incidences et leur pondération est aussi effectuée dans d'autres plateformes internationales [152, 153].

Les données issues des cohortes participatives soulèvent néanmoins de nouvelles questions concernant l'analyse des données, en raison de leur nature non traditionnelle. Cela concerne la participation dynamique des individus, avec entre autres des comportements de déclaration variables au cours du suivi, mais aussi l'interprétation et la validation des données recueillies.

Concernant la participation, sa prise en compte est importante dans le calcul des incidences, car elle impacte l'estimation du signal. Plusieurs études ont montré que beaucoup de participants ne prennent part au suivi participatif que pour déclarer des symptômes [138, 144, 146, 148, 149, 163, 164]. Ce phénomène est d'ailleurs particulièrement observé lors de la toute première participation à l'étude, avec la déclaration presque systématique de symptômes [144, 146, 149, 163]. On retrouve cette tendance dans d'autres systèmes participatifs comme Flu Near You aux USA, ou GrippeWeb en Allemagne [152, 153]. La déclaration systématique de symptômes engendre alors une surestimation des épisodes observés, et donc de l'incidence. Afin de réduire ce biais, la définition de participants « actifs » a été avancée au sein d'InfluenzaNet, pour ne prendre en compte que les participants qui rapportent un état de santé régulièrement, mais aussi en excluant des analyses le premier questionnaire de tous les nouveaux participants [144, 149, 163]. Ces critères de prise en compte de la participation sont des hypothèses avancées, propres aux systèmes participatifs, mais n'ont pas fait l'objet d'une évaluation systématique et globale.

En parallèle, la détermination des cas s'est souvent focalisée sur une définition de cas unique, celle du syndrome grippal de l'ECDC [144, 149]. La valeur ajoutée de la flexibilité du système n'est donc pas véritablement étudiée, puisque l'on ne sait pas comment d'autres combinaisons de symptômes auto-déclarés peuvent impacter l'estimation des incidences des épisodes grippaux, et quelles sont les différences entre les définitions de syndrome grippal et d'IRA à partir de ce système participatif. Pour éviter de reconstruire des définitions de cas, un travail récent mené au sein des différents pays d'InfluenzaNet, a effectué l'étude inverse en partant directement des symptômes déclarés, sans imposer de définition de cas a priori. La justification de cette étude part du constat qu'il n'y a pas de définition consensuelle d'un épisode grippal, et que les épidémies de grippe saisonnière sont différentes entre les pays. L'objectif était de voir si cette approche est capable de fournir un signal épidémiologique valide qui ne repose pas sur une définition de cas en le comparant aux systèmes sentinelles traditionnels [165].

Malgré une population non-représentative en termes d'âge et de sexe, les tendances des incidences des syndromes grippaux observées sont comparables à celles issues des systèmes de surveillance traditionnelle des différents pays [166]. Cela est aussi retrouvé avec le travail européen sur l'approche non supervisée des symptômes [165]. Le pic épidémique est souvent anticipé d'une semaine [138, 144, 146, 149, 163]. Ce résultat est d'ailleurs identifié dans d'autres pays [152, 153], suggérant que le temps nécessaire pour consulter un médecin généraliste qui remonte ensuite l'information au système de surveillance serait d'une semaine. *Kjelso et al.* expliquent dans le cas de la plateforme danoise, que l'envoi du questionnaire hebdomadaire effectué un jour donné (J) génère 90% de l'ensemble des données reçues à J+1, ce qui n'est pas le cas des systèmes traditionnels [149]. Ces mêmes tendances ont été observées dans les autres pays européens, dont la France.

Finalement, même si les premières estimations des incidences des signaux grippaux sont encourageantes au sein de la plateforme InfluenzaNet, il manque une analyse systématique pour comprendre le système dans sa globalité : impact de la participation et de la déclaration des épisodes par les participants, couplées aux signaux obtenus par le choix de la définition de cas de l'épisode grippal (syndrome grippal vs. IRA). Le travail européen InfluenzaNet a montré la flexibilité du système dans le suivi des signaux grippaux directement à partir des symptômes collectés. Cependant, aucune étude n'a comparé le signal participatif avec des signaux de syndrome grippal et d'IRA issus de la surveillance traditionnelle dans un même pays. Il est important de pouvoir étudier de façon globale tous ces critères pour pouvoir valider ce système de surveillance à partir de la population générale. C'est l'objectif de notre premier travail de thèse, qui sera détaillé dans le Chapitre 4.

3.2.4. Autres travaux de recherche menés par le biais de la surveillance participative des syndromes grippaux en Europe

Pour caractériser au mieux les épidémies de grippe directement à partir de la population, les données épidémiologiques collectées via la plateforme InfluenzaNet ont permis la mise en place d'études sur l'estimation de l'efficacité vaccinale antigrippale, le recours aux soins et sur les facteurs associés aux syndromes grippaux.

L'**efficacité vaccinale antigrippale** est estimée chaque année, afin d'alerter le plus précocement possible les autorités de santé dans le but de mettre en place des actions de prévention ciblées (cf. personnes à risque de complications), au cas où le vaccin s'avère inefficace face aux virus grippaux circulants. L'estimation de cette efficacité est réalisée à partir du système de soins. En effet, les cas doivent être identifiés cliniquement ou confirmés virologiquement. Cependant, des différences peuvent être observées dans la présentation des symptômes entre les individus vaccinés et non vaccinés, entraînant des différences potentielles dans le dépistage par les professionnels de santé. Au vu des informations collectées à travers la surveillance participative : présence d'un syndrome grippal ou non pour une semaine donnée, statut vaccinal antigrippal, âge, groupes à risque de complication, ce nouveau système est un bon candidat pour estimer l'efficacité vaccinale en dehors du système de soins en temps réel. Les équipes britanniques et françaises ont mené des études sur l'efficacité vaccinale lors des saisons grippales 2010/2011 et 2012/2013 respectivement [167, 168]. Les deux études ont mis en évidence la possibilité d'estimer une efficacité vaccinale rapidement et directement à partir de la population. Les estimations réalisées étaient concordantes avec celles issues des systèmes traditionnels. Dans l'étude française, l'efficacité vaccinale a été estimée de façon rétrospective (une fois toutes les données collectées à la fin de la saison) et en temps réel. Les deux analyses ont montré des résultats probants, même si l'estimation en fin de saison s'est avérée plus précise. La valeur ajoutée de ces estimations en population générale peut être particulièrement utile dans un contexte d'urgence comme le contexte pandémique, lorsque le système de soins est engorgé.

Le **recours aux soins** en cas de syndrome grippal a été étudié au sein d'InfluenzaNet, dans la plupart des pays partenaires : Belgique, Espagne, France, Italie, Pays-Bas, Portugal, Royaume-Uni, Suède [143, 166, 169]. Cette estimation, ainsi que la connaissance des déterminants liés à la consultation d'un professionnel de santé en cas de syndrome grippal, paraît essentielle pour caractériser l'accessibilité aux soins et mettre en place des actions de prévention ciblées, comme chez les personnes à risque de complications liées à la grippe. Le recours aux soins variait beaucoup selon les pays, allant de moins de 25% en Suède à près de 75% en Belgique. En Italie, 36% avait recours à un médecin généraliste en cas de syndrome grippal, et ce pourcentage était plus important chez les enfants [143]. En France, 33% et 57% des participants avaient recours au médecin généraliste lors d'une IRA (définition de l'ECDC) et d'un syndrome grippal (définition du réseau Sentinelles) respectivement [169]. Ces chiffres étaient similaires à ceux retrouvés par d'autres moyens de surveillance [16, 170]. Les différences observées entre les pays sont

retrouvées également lorsque l'on compare le délai entre le début des symptômes et la visite chez le médecin généraliste [166]. Le recours au système de soins en cas d'épisode grippal semble lié à des différences sociétales et économiques entre les pays. Les déterminants du recours au système de soins en France sont : vivre en zone rurale ; avoir un niveau d'éducation peu élevé ; utiliser Internet pour chercher des informations sur la grippe ; présenter des symptômes inquiétants comme la fièvre, la toux, une dyspnée, une asthénie ou des expectorations ; percevoir négativement son état de santé ; avoir déclaré un médecin traitant à l'Assurance maladie. Les participants qui utilisaient des médecines alternatives pour se soigner se rendaient moins fréquemment chez le médecin généraliste [169]. Il n'y avait pas de différence de recours aux soins entre les participants et ceux étant à risque de complications.

La recherche des **facteurs associés aux syndromes grippaux** est importante pour cibler les personnes les plus à risque d'infections avec des actions de prévention. Habituellement, ces études sont conduites via le système de soins et notamment en milieu hospitalier pour connaître les facteurs de complication dus à la grippe. Grâce à la surveillance participative et à ses données épidémiologiques fines, la connaissance de ces facteurs pourrait permettre d'éviter des cas plus graves, en limitant le risque de transmission. Différents travaux ont été menés par le biais d'InfluenzaNet, à l'échelle européenne [139, 166] ou nationale [144, 171]. Des facteurs similaires ont été retrouvés à travers les différents pays, et concordaient avec ceux retrouvés dans la littérature [139]. Le détail de l'analyse des facteurs associés aux syndromes grippaux fera l'objet d'une discussion plus spécifique dans le Chapitre 5, où nous présenterons le second travail réalisé dans cette thèse. Cette étude a permis de compléter les analyses faites au niveau européen, en les menant sur six saisons en France pour avoir des estimations plus robustes des facteurs de risque de syndrome grippal. Pour préciser ces déterminants, elles ont également été conduites à partir des deux indicateurs grippaux : IRA et syndrome grippal, ce qui n'avait pas encore été évalué.

Même si la plateforme commune InfluenzaNet promeut une certaine homogénéité dans la collecte et l'analyse des données entre les différents pays, chacun des pays continue à avoir des projets scientifiques qui lui sont propres. Ceci est particulièrement vrai pour la plateforme française qui met en place différents projets de surveillance et de recherche à l'échelle européenne et nationale. Des questions ou questionnaires peuvent alors être ajoutés pour collecter de nouvelles données et mieux suivre et comprendre la grippe directement à partir de ces cohortes participatives. Nous allons détailler ci-après la cohorte GrippeNet.fr et son application dans le champ de la surveillance en population générale en France.

3.3. La surveillance participative en France avec GrippeNet.fr

Depuis 2012, le dispositif de surveillance GrippeNet.fr est mis en place en France métropolitaine par le réseau Sentinelles, sous la tutelle de l'Inserm et de Sorbonne Université, et par SpF. Le projet a été

initialement financé par l'Agence Nationale de la Recherche jusqu'à l'été 2016, et est financé depuis par SpF. Aucun financement privé n'a été reçu pour développer le projet.

Intégré au consortium européen InfluenzaNet, il est basé sur la collecte d'informations épidémiologiques sur la grippe directement à partir de la population générale. Le détail du fonctionnement de ce système participatif a été décrit précédemment (§ 3.2.1). L'intérêt de cette surveillance est de compléter la surveillance traditionnelle de la grippe, en collectant des données auprès de personnes qui ne consultent pas forcément le système de soins en cas d'épisode grippal, pour avoir une vision globale de l'impact des épidémies de grippe. Le questionnaire d'inclusion avec la collecte des caractéristiques des participants et le questionnaire hebdomadaire de symptômes sont disponibles respectivement en Annexe 1 et 2.

A travers cette section, nous allons décrire le système français GrippeNet.fr en présentant d'abord la participation et le recrutement de la population d'étude, ainsi que le retour d'information effectué. Grâce aux données épidémiologiques fines collectées à un niveau individuel, différentes études ont pu être menées pour mieux comprendre l'impact de la grippe dans la population. Nous détaillerons notamment les travaux réalisés sur le suivi de groupes de population ciblés par les recommandations vaccinales (femmes enceintes, personnes immunodéprimées), sur l'opinion de la population concernant la vaccination antigrippale, et sur les connaissances et comportements des participants face à la grippe. Nous présenterons dans le Chapitre 6 l'usage fait de la cohorte GrippeNet.fr au-delà de la grippe, avec (i) le suivi d'autres pathologies infectieuses comme les gastroentérites aiguës, ou (ii) d'autres sujets de santé publique pour lesquels le recours à un système participatif peut être avantageux.

3.3.1. Recrutement, participation et communication

Après une première saison pilote lancée en janvier 2012, la **participation** à GrippeNet.fr s'est très vite stabilisée autour 6 000 participants par saison (Figure 8a). Il s'agit de participants qui ont rempli au moins un questionnaire d'inclusion lors d'une saison donnée. Les participants représentent environ 0,01% de la population française métropolitaine. Chaque année, environ 1 000 nouveaux participants intègrent GrippeNet.fr. Cependant, 80% des GrippeNautes d'une saison donnée sont des anciens participants. Beaucoup de participants sont donc fidèles à la cohorte. Lors de la dernière saison 2018/2019, 25% de l'ensemble des GrippeNautes étaient présents depuis le lancement en 2012, soit depuis 8 saisons de surveillance (Figure 8b). Concernant la participation hebdomadaire, celle-ci comptabilise chaque année environ 4 000 participants tout au long de la saison. Lors de la dernière saison 2018/2019, 4 226 questionnaires ont été renseignés en moyenne par semaine (Figure 8c). La participation étant volontaire, chaque personne inscrite peut remplir les questionnaires hebdomadaires à son rythme durant la saison hivernale. Les déterminants de la participation ont été analysés au niveau européen en 2014. Les données de GrippeNet.fr ont été analysées avec celles de la Suède, du Royaume-Uni, des Pays-Bas, de la Belgique, de l'Italie et du Portugal. *Bajardi et al.* ont montré que la participation étaient liée à l'éducation et au mode de vie [172]. Les fumeurs, les plus jeunes, les personnes avec un

niveau d'éducation peu élevé, celles vivant avec des enfants ou encore celles qui ne s'étaient pas fait vacciner contre la grippe avaient une participation moindre. Des différences pouvaient avoir lieu entre les pays, mais les résultats demeuraient similaires. Cela se vérifiait donc pour la France, avec par ailleurs une meilleure participation des femmes par rapport aux hommes [172]. Une étude similaire conduite à partir de la cohorte participative Flu Near You aux USA a mis en évidence les mêmes déterminants que ceux retrouvés au sein de notre plateforme [173].

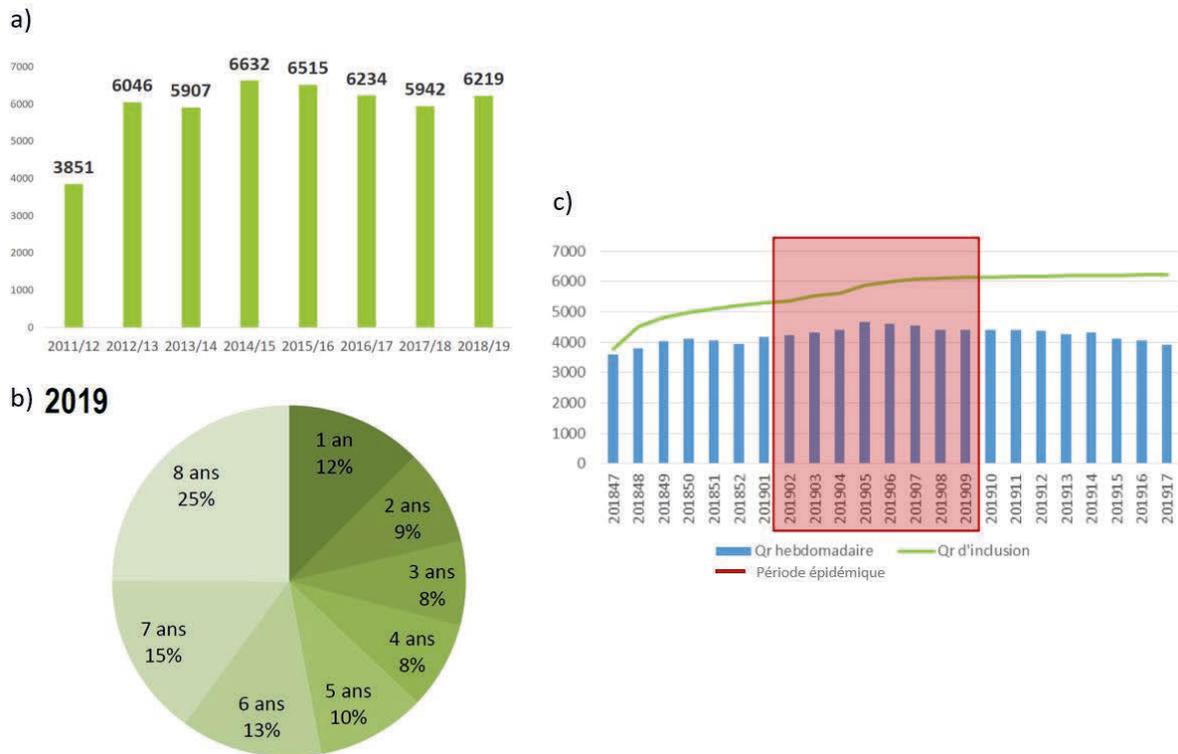
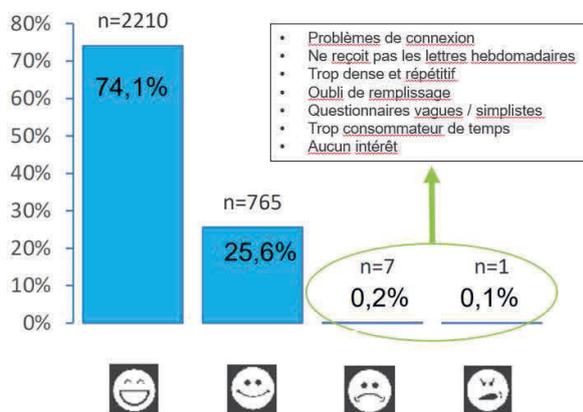


Figure 8. Participation à GrippeNet.fr. (a) Nombre de participants par saison (personnes ayant rempli au moins un questionnaire d'inclusion) ; (b) Ancienneté des participants en 2019 ; (c) Participation hebdomadaire en bleu et évolution du nombre de questionnaires d'inclusion remplis en vert.

Pour mieux répondre aux attentes des GrippeNautes, deux enquêtes de satisfactions ont été réalisées en [2013/2014](#) et en [2015/2016](#). A chaque fois, un taux de réponse de près de 50% était observé, soit environ 3 000 réponses obtenues. Les réponses étaient similaires en 2013/2014 et 2015/2016. La quasi-totalité des participants répondants étaient satisfaits de leur participation (99,7%) et presque tous souhaitaient participer l'année suivante (98,6%) (Figure 9a). Plus de 90% trouvaient les questionnaires d'inclusion et hebdomadaire de la bonne longueur, et plus de 85% étaient prêt à répondre à des questionnaires supplémentaires posés ponctuellement pendant ou en dehors de la saison grippale, quelle que soit la thématique de santé concernée (Figure 9b). Les GrippeNautes étaient confiants dans la capacité à évaluer leurs symptômes (89%).

a)

Globalement êtes-vous satisfait(e) de GrippeNet.fr ?



b)

Longueur des questionnaires

	Bonne longueur	Trop long	Trop court	NA
Qr. d'inclusion	96,4%	1,6%	0,6%	1,3 %
Qr. hebdomadaire	93,4%	3,8%	1,6%	1,2%

Extra-questionnaires

	Oui, peu importe le sujet	Oui, seulement si cela concerne la grippe	Non	Je ne sais pas
Pendant la saison grippale	88,6%	5,3%	1,1%	5,0%
En dehors de la saison grippale	86,2%	3,1%	3,4%	7,3%

Figure 9. Satisfaction des participants lors de la saison 2015/2016. a) Satisfaction générale ; b) Satisfaction par rapport aux questionnaires proposés

Le **recrutement** des participants est effectué par l'équipe en charge du développement du projet. Depuis sa mise en place, différentes actions de communication ont été conduites. Elles ont lieu surtout à l'automne, avant le lancement de l'étude, puis ponctuellement tout au long de la saison. Les campagnes de recrutement sont d'abord portées par les tutelles et s'effectuent avant tout sur Internet. Le réseau Sentinelles, l'Inserm, Sorbonne Université et SpF relaient sur leur site Internet respectif des appels à participation. C'est surtout ce mode de communication qui permet l'adhésion du plus grand nombre de participants. En effet, plus de 60% des 6 000 participants sont recrutés par ce biais. Le bouche-à-oreille est également un bon moyen de recrutement. Lors de la dernière saison 2018/2019, près de 40% des GrippeNautes avaient connu le projet via leur famille, amis, collègues, professionnels de santé ou par le système email de parrainage mis en place. La communication s'est faite également par des campagnes d'affichage dans des pharmacies et la diffusion de vidéo dans les salles d'attente de cabinets médicaux. Une étude réalisée en 2014 au sein de la plateforme InfluenzaNet a montré que les campagnes de recrutement sur Internet pour le suivi de cohortes participatives étaient moins efficaces que le recrutement de participants hors ligne. Seul GrippeNet.fr faisait exception, avec des campagnes de communication pour le recrutement en ligne plus efficaces que celles réalisées hors ligne (via les affiches distribuées) [174].

Le **retour d'information** et la **communication** réalisés à partir des données collectées se fait au sein de la cohorte GrippeNet.fr, mais aussi en dehors grâce à leur valorisation. Dans le cadre interne, les participants reçoivent chaque semaine entre novembre et avril une lettre d'information hebdomadaire dans laquelle ils peuvent retrouver des résultats préliminaires sur la situation des syndromes grippaux au sein de la cohorte, et sur la situation mesurée par le système de surveillance traditionnelle (Figure 10a). D'après la dernière enquête de satisfaction, 58% lisaient la lettre toutes les semaines ou assez souvent. Cette lettre hebdomadaire semble plutôt

importante pour promouvoir la participation des GrippeNautes, puisque 90% ont exprimé se rendre sur le site Internet GrippeNet.fr uniquement quand ils remplissaient le questionnaire hebdomadaire de symptômes. Par ailleurs, les participants reçoivent aussi des retours spécifiques sur les études menées grâce aux données recueillies et des lettres estivales retraçant le bilan de la saison ou détaillant d'autres sujets de santé. En plus des lettres envoyées par mail, des résultats plus ciblés sont disponibles sur leur espace personnel GrippeNet.fr et sur le site Internet www.grippenet.fr (Figure 10b). Cependant, les participants semblent plus sensibles à l'information que nous pouvons leur apporter, que par le fait de rechercher eux-mêmes les informations sur nos médias. En effet, seul 21% visite le site Internet GrippeNet.fr et 13% utilisent nos réseaux sociaux.

Concernant la valorisation des données issues de GrippeNet.fr à l'extérieur de la cohorte, celles-ci sont reprises depuis 2012 dans le bulletin épidémiologique hebdomadaire du réseau Sentinelles (Figure 10c). Ce soutien des acteurs de la surveillance traditionnelle en France est primordial, car il permet d'appuyer l'importance de GrippeNet.fr en complément de la surveillance effectuée par les médecins généralistes aux yeux du grand public, mais aussi auprès des instances sanitaires. Plus largement, des discussions sont en cours avec l'ECDC pour intégrer les données InfluenzaNet dans leur bulletin grippe hebdomadaire (i.e. Weekly Influenza update). Cet intérêt porté par les autorités sanitaires permettra d'ancrer véritablement la surveillance participative dans le champ de la surveillance sanitaire, et aidera à la promotion de son intégration dans les dispositifs nationaux de surveillance de la grippe.

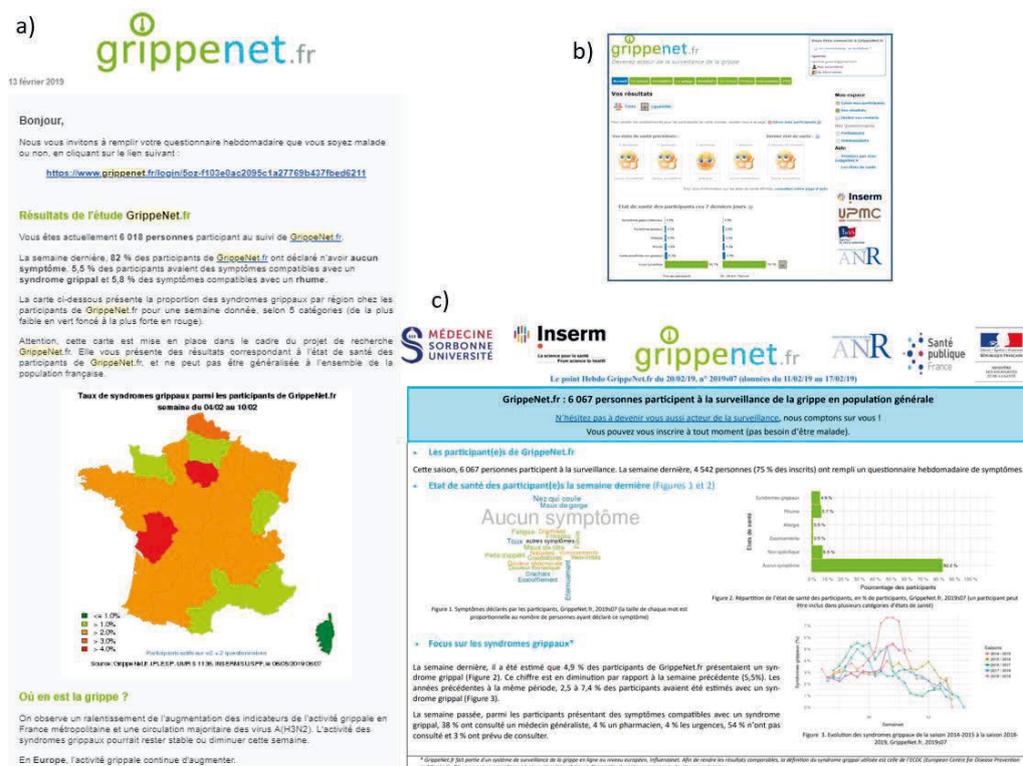


Figure 10. Communications réalisées pour la diffusion des résultats issus de l'analyse des données collectées par le biais de GrippeNet.fr. (a) Lettre d'information hebdomadaire ; (b) espace personnel des participants ; (c) intégration des données GrippeNet.fr au bulletin épidémiologique du réseau Sentinelles.

3.3.2. GrippeNet.fr pour mieux comprendre la grippe

Hormis les travaux réalisés dans le cadre de ce travail de doctorat pour évaluer la surveillance de la grippe en population, avec l'estimation des incidences des syndromes grippaux issues de GrippeNet.fr (Chapitre 3) et la détermination des facteurs associés aux syndromes grippaux (Chapitre 4), d'autres études ancillaires ont été menées pour mieux comprendre la grippe à cette échelle.

3.3.2.1. Evaluer l'opinion des GrippeNautés sur la vaccination antigrippale

Dans un contexte de défiance vaccinal croissant, la gestion de la campagne de vaccination en France lors de la pandémie grippale de 2009 semble avoir particulièrement impacté l'acceptabilité de la vaccination dans la population, et en particulier celle à l'égard de la vaccination antigrippale saisonnière [175, 176]. La France représente d'ailleurs l'un des pays où le scepticisme vaccinal est le plus important au monde [177]. Dans ce contexte, *Boiron et al.* ont mené une étude à partir de GrippeNet.fr afin d'évaluer l'acceptabilité du vaccin contre la grippe saisonnière dans la population générale et de déterminer les facteurs associés à cette acceptabilité, trois ans après la pandémie grippale de 2009. La connaissance des déterminants de l'acceptabilité de la vaccination antigrippale par la population peut être utile aux autorités sanitaires, qui cherchent d'ailleurs à faciliter l'accès à la vaccination avec l'ouverture aux pharmaciens de la vaccination antigrippale depuis 2017. Parmi les 5 374 participants inclus, 39% rapportaient une opinion positive, 39% une opinion neutre et 22% une opinion négative. Les facteurs associés à une opinion neutre ou négative étaient : un jeune âge, un niveau d'éducation peu élevé, une absence de contact avec des personnes âgées ou malades, une absence de traitement pour une de pathologie chronique sous-jacente, et la prise de traitements homéopathiques préventifs pour lutter contre la grippe [178]. Cette étude a permis de confirmer la hausse des personnes défavorables au vaccin antigrippal en France lors de la saison 2012/2013. Elle a permis d'esquisser le profil des personnes particulièrement réticentes qui pourraient être la cible des campagnes d'information à mettre en place par les autorités de santé.

3.3.2.2. Suivre des sous-cohortes particulières

Comme la grippe peut avoir un impact sanitaire important parmi les personnes à risque de complications, un suivi spécifique de ces populations directement à partir de la population paraît opportun pour mieux connaître l'impact des syndromes grippaux et de certains déterminants. Depuis la mise en place de GrippeNet.fr, deux sous-cohortes ont été construites. La première a permis le suivi des femmes enceintes et la seconde (en cours d'analyse) le suivi de personnes immunodéprimées.

Depuis 2012, **les femmes enceintes** sont ciblées par les recommandations vaccinales antigrippales en France. Cependant, peu d'informations sont connues dans cette population spécifique comme l'incidence des syndromes grippaux. Grâce à GrippeNet.fr, le suivi de 153 femmes enceintes a été effectué lors de la saison 2014/2015, et s'est poursuivi jusqu'à la saison 2016/2017. Par rapport aux femmes enceintes en France, les femmes de la cohorte étaient plus âgées, avec un niveau d'éducation plus élevé. Cependant, les mêmes caractéristiques étaient retrouvées concernant certains indicateurs liés à la grossesse comme la parité. Les femmes semblaient adhérer à ce type de suivi participatif puisque le taux médian de participation était de 78%. Une participation active à GrippeNet.fr était observée plus fréquemment chez les femmes avec un niveau d'éducation élevé et parmi celles qui avaient participé à une saison antérieure [179]. Le taux d'incidence des syndromes grippaux calculé sur l'ensemble de la période de suivi chez ces femmes enceintes était de 16% (IC95% [11 ; 23]), et la densité d'incidence de 1,8 / 100 personnes-semaine (IC95% [1,1 ; 2,6]). L'incidence et la densité d'incidence ne différaient pas par rapport aux autres femmes en âge de procréer suivies via GrippeNet.fr, respectivement 14% (IC95% [13 ; 16], $p = 0,26$) et 1,3 / 100 personnes-semaine (IC95% [1,2 ; 1,6], $p = 0,23$). La densité d'incidence était plus élevée chez les femmes de 40 ans et plus, et lors des premier et deuxième trimestres par rapport au troisième trimestre. Concernant la couverture vaccinale antigrippale, celle-ci était estimée à 39% (IC95% [31 ; 47]) et à 26% (IC95% [20 ; 34]) après ajustement sur l'âge et le niveau d'éducation. Les facteurs associés à une vaccination accrue étaient : une recommandation par un professionnel de santé ; être non fumeuse [180]. Les deux travaux menés à partir de cette sous-cohorte ont mis en évidence notamment (i) les caractéristiques des femmes enceintes qui ont une incidence de syndrome grippal plus élevée et (ii) la faible valeur de la couverture vaccinale, dans une population pour laquelle la vaccination antigrippale est recommandée.

Sur le même principe, le suivi d'une cohorte de **personnes immunodéprimées** a été effectué lors de la saison 2017/2018. L'objectif était d'estimer au sein de cette cohorte issue d'un échantillon de patients traités par immunosuppresseurs, biothérapie et/ou corticothérapie pour une maladie auto-immune ou inflammatoire chronique : l'incidence des syndromes grippaux lors de l'hiver 2017/2018, et les facteurs associés à la vaccination antigrippale, car peu d'information est disponible. Les analyses sont actuellement en cours. Cette cohorte a permis le suivi de 177 personnes. Là encore, les participants ont adhéré au système, avec un taux de participation médian sur la période de suivi de 65% (IC95% [22% ; 87%]).

3.4. Avantages et limites de la surveillance participative

A travers les différents systèmes de surveillance participatifs en place dans le monde, les valeurs ajoutées les plus souvent citées sont : la possibilité de suivre les événements de santé en temps réel et de façon rapide [129, 133, 139], et la flexibilité permettant le suivi de différents indicateurs de santé au sein d'un même système de surveillance, avec l'ajustement de ces indicateurs en temps réel [136, 139]. Sur le plan économique, la mise en place et la gestion de ces outils

participatifs représentent des coûts peu élevés et fixes peu importe le nombre de participants intégrés [129, 133, 136, 139]. Par ailleurs, la collecte épidémiologique et le développement technique des plateformes Internet peuvent être standardisés à travers différents pays, comme dans le cadre du consortium InfluenzaNet. Cela permet d'un côté de standardiser l'information collectée à une échelle internationale, mais aussi de mutualiser les coûts inhérents au développement de ces outils participatifs. Sur le plan sanitaire, la sensibilité de ces systèmes permet le suivi d'événements inhabituels souvent d'origine infectieuse, mais aussi le suivi de pathologies non infectieuses. Enfin, sur le plan épidémiologique, ces systèmes participatifs permettent d'accroître la couverture de la population intégrée aux systèmes de surveillance [87]. Grâce au niveau des données collectées et à leur précision, des analyses épidémiologiques fines peuvent être conduites pour mieux connaître les événements de santé surveillés [133, 136, 139]. Par rapport aux autres sources de données digitales, les systèmes participatifs sont plus précis dans la nature des données collectées et indépendants des autres sources.

Les limites de la surveillance participative apparaissent surtout dans la nature des données collectées puisqu'elles sont recueillies directement à partir d'une population non experte, contrairement à la surveillance traditionnelle. Les critiques concernent donc le manque de validité, de spécificité et la fiabilité de l'information collectée [133, 137, 173, 181]. Les différentes études menées en France, en Europe ou dans le monde ont montré que les signaux obtenus étaient bien corrélés aux systèmes traditionnels, prouvant la possibilité du système à suivre les signaux grippaux, voire à anticiper le signal émis. Dans cette même lignée, l'absence de confirmation virologique est alors souvent reprochée [136]. Cela ne semble donc pas être un problème, d'autant que deux études ont démontré la faisabilité de la mise en place d'une surveillance virologique des syndromes grippaux à partir d'auto-prélèvements nasopharyngés réalisés directement via la population [182, 183]. Par ailleurs, les participants pourraient aussi être influencés par des facteurs externes, comme les médias, dans le rapport des symptômes présentés [181]. Jusqu'à présent, seule la primo-participation semble surestimer le signal obtenu car les participants vont plus facilement déclarer des symptômes. Cependant, cela peut être corrigé. Seul un suivi au long-courant permettra de connaître l'impact des facteurs externes sur le signal émis par ces systèmes participatifs.

Comme pour la surveillance syndromique digitale (avec les données d'Internet, réseaux sociaux), les données collectées peuvent ne pas être représentatives de la population générale, et cela d'autant plus que les systèmes participatifs sont essentiellement basés sur la participation volontaire des individus [181]. La recherche de la représentativité est l'un des principes fondamentaux dans la mise en place d'un système de surveillance [1]. Ce problème n'est donc pas spécifique à la surveillance participative, mais s'applique aussi à la surveillance traditionnelle qui peut rencontrer les mêmes obstacles. Le recrutement et le maintien de la participation des individus représentent donc un challenge majeur [136, 181]. Dans le cadre de GrippeNet.fr, les efforts réalisés en matière de communication interne et externe semblent aider à la pérennisation de la cohorte et à la promotion de la participation des GrippeNautes aux différentes

études proposées. En effet, les taux de participation sont souvent supérieurs ou égal à 50% de l'ensemble des participants inscrits.

Enfin, la nouveauté de ces sources d'informations liées à la santé peut soulever des questions plus larges sur le traitement de ce type de données personnelles, et donc l'aspect éthique de l'utilisation de ces données. Un cadre éthique permet d'une part d'assurer la confiance et l'engagement des participants pour permettre la durabilité et la qualité de ces systèmes de surveillance participatifs. Si les participants percevaient un risque de violation de la vie privée, ils pourraient s'abstenir de renseigner des informations nécessaires à une surveillance efficace [184]. Le cadre éthique proposé par *Geneviève et al.* et directement appliqué à InfluenzaNet comporte quatre composants essentiels : avoir un consentement électronique, protéger la vie privée des participants, donner accès à l'information sur la surveillance effectuée et aux stratégies de prévention, faire valider le système par un Comité d'éthique [159]. Ces composants sont déjà mis en place dans la cohorte GrippeNet.fr depuis son lancement en 2012.

Finalement, pour pallier les limites des systèmes participatifs, l'important est de bien connaître le système en question et la nature des données collectées. Cela passe alors par la mise en place de différentes études épidémiologiques nécessaires à la connaissance des caractéristiques des participants, et à l'analyse et à l'interprétation des signaux captés.

Les messages clés du Chapitre 3

- La surveillance participative permet la collecte active de données de santé par le biais de la population générale volontaire et via Internet, afin de comprendre la dynamique des maladies.
- Les cohortes participatives permettent des activités de surveillance et de recherche en santé publique, grâce à la collecte de données épidémiologiques fines.
- La flexibilité des systèmes participatifs permet le suivi de différents indicateurs grippaux grâce à la collecte de symptômes identifiables facilement.
- Par définition, la population participant à cette surveillance n'est pas représentative de la population générale, puisqu'elle est volontaire. Des ajustements sont possibles. La surveillance participative permet de capter la population qui ne consulte pas le système de soins.
- Les tendances observées à partir de la surveillance participative sont similaires à celles des systèmes traditionnels. Le pic épidémique est souvent anticipé.
- Les participants des cohortes participatives sont attentifs à leur santé. Ils sont plutôt fidèles dans leur participation. L'envoi de lettre d'information hebdomadaire, pour collecter et retourner l'information, est le principal moteur de la participation. Le soutien des institutions sanitaires nationales est primordial pour le développement et la pérennisation de ces cohortes.
- L'enjeu principal consiste à bien connaître le système pour pouvoir le valider et le promouvoir auprès des autorités sanitaires.

Chapitre 4.

L'intérêt potentiel des systèmes de surveillance participatifs pour compléter les réseaux sentinelles de la grippe : le cas de la France

L'évolution de l'épidémiologie traditionnelle vers une épidémiologie digitale ouvre une voie facilitée vers l'usage d'autres sources d'information dans les domaines de la recherche et de la surveillance en santé. Dans le cadre de la veille sanitaire, la surveillance réalisée directement par le biais de la population représentait le dernier niveau encore non investi en France. La surveillance participative à travers GrippeNet.fr permet alors le recueil de données en lien avec la grippe de façon volontaire pour mieux caractériser les épidémies de grippe. Ce nouveau niveau de surveillance permet la collecte de données plus sensibles que celles obtenues via le système de soins. Cependant, une compréhension globale du dispositif est nécessaire à l'estimation et à l'interprétation des données collectées concernant les syndromes grippaux. Dans le cadre de la surveillance de la grippe en population générale, la définition d'une méthode de calcul des incidences propre à GrippeNet.fr et l'interprétation des résultats obtenus sont primordiales. Elles permettront de valider ce dispositif participatif dans sa capacité à suivre les épidémies de grippe en complément des systèmes traditionnels.

4.1. Problématiques liées à l'estimation des incidences des syndromes grippaux

L'estimation de la morbidité liée à la grippe s'effectue de différentes façons selon les dispositifs de surveillance. Au niveau hospitalier, le réseau OSCOUR® enregistre le nombre de passages aux urgences et le nombre d'hospitalisations après passage aux urgences pour syndrome grippal. Au niveau des soins primaires, SOS Médecins rapporte la proportion de visites pour syndrome grippal parmi l'ensemble des actes médicaux renseignés par les professionnels de santé participants. Si ces méthodes d'estimation nécessitent simplement la connaissance du nombre total de passages, d'hospitalisations ou d'actes médicaux, elles ne permettent pas de quantifier la fréquence de la maladie dans la population puisque les populations cibles sont spécifiques. Par ailleurs, pour OSCOUR®, les indicateurs suivis sont liés à l'évolution continue du nombre d'établissements participants, contraignant le système à les calculer à partir d'un nombre d'établissements constant pour mesurer les changements de tendance [185]. Concernant le réseau Sentinelles,

l'indicateur de santé utilisé est le taux d'incidence des consultations pour syndrome grippal pour 100 000 habitants [14]. Comme détaillé dans le Chapitre 1 (§1.2.5), cet estimateur repose sur deux hypothèses principales qui sont que (i) les médecins participants sont représentatifs de l'ensemble des médecins généralistes de France métropolitaine, et que (ii) les déclarations qu'ils effectuent sont représentatives de leur activité hebdomadaire. En effet, la taille de la population couverte par les médecins généralistes n'est pas connue en France, car il n'existe pas de registre de pratique obligatoire [14]. Par ailleurs, l'estimation des données finales est obtenue après une période de consolidation de trois semaines, comme les médecins peuvent déclarer des cas remontant à 12 jours. Le délai de consolidation en situation épidémique peut donc représenter un frein dans la gestion de l'épidémie par les autorités sanitaires en temps réel. Enfin, ces dispositifs de surveillance hospitaliers ou en soins primaires peuvent être saturés en cas d'afflux exceptionnel de la population, comme ce fut le cas lors de la pandémie grippale de 2009. L'estimation précise de la situation par le biais de ces dispositifs peut s'avérer difficile voire impossible.

L'intérêt principal du système participatif GrippeNet.fr est sa flexibilité dans le recueil des données. Actuellement, celui-ci est calibré sur un recueil hebdomadaire. Avant le travail présenté ci-après, l'estimation des syndromes grippaux reposait semaine après semaine sur le rapport entre le nombre de participants ayant déclaré des symptômes compatibles avec un syndrome grippal (numérateur) et le nombre total de participants pour une semaine donnée (dénominateur). Les problèmes fondamentaux liés à cette méthode d'estimation préliminaire sont que :

- La participation hétérogène des participants n'est pas prise en compte.
- Le comportement des participants dans le rapport et l'estimation de leur état de santé est inconnu.
- Les résultats ne sont pas directement transposables à la population générale résidant en France métropolitaine, puisque les participants ne sont pas représentatifs et qu'aucune méthode de redressement des données n'était effectuée jusqu'alors.

Ainsi, nous avons voulu étudier et comprendre l'impact du numérateur et du dénominateur dans les calculs d'estimation des incidences à partir de ce nouveau dispositif. Concernant l'étude du numérateur, les travaux menés au niveau européen n'ont pas permis d'exploiter la flexibilité du système et de comprendre l'impact des symptômes décrits par les participants sur l'estimation du signal grippal. Le système est-il assez sensible pour détecter un signal tout en étant suffisamment spécifique pour suivre les syndromes grippaux ? Le présent travail s'est donc attaché à étudier la capacité de la cohorte à évaluer son état de santé à partir des symptômes recueillis. En effet, l'interprétation de l'état de santé peut être différente entre les professionnels et la population, et donc impacter la mesure des signaux de l'activité grippale effectuée. Il est essentiel de comprendre spécifiquement l'interprétation faite par les participants pour identifier et mettre en place une définition de cas optimale à ce nouveau dispositif de surveillance, indépendamment des définitions utilisées dans la surveillance traditionnelle. L'étude du dénominateur vise à comprendre et prendre en compte le comportement des participants et leur

participation hétérogène dans le dispositif. L'objectif global est finalement de déterminer le meilleur indicateur de suivi des syndromes grippaux propre à ce système participatif.

L'estimation des indicateurs liés à la grippe représente un enjeu majeur pour le suivi des épidémies hivernales. A tous les niveaux de surveillance, les dispositifs de collecte de données épidémiologiques possèdent des contraintes qui leur sont propres. Les méthodes d'analyse doivent leur être adaptées. L'amélioration de ces dispositifs dans la qualité des signaux produits reste cependant continue (i.e. exhaustivité, représentativité).

4.2. Résumé de l'article

La surveillance participative fournit des données épidémiologiques directement à partir de la population. Ces données nous permettent de connaître leurs caractéristiques et état de santé durant le suivi. Ce nouveau système de surveillance apporte aussi de nouvelles problématiques et questions qu'il est important d'étudier pour comprendre son fonctionnement et juger de son utilité dans le dispositif de surveillance établi.

Dans ce premier travail, nous avons exploré l'utilité des données de GrippeNet.fr pour estimer l'incidence des maladies liées à la grippe en France en population. Depuis 2012, près de 5 000 participants par semaine signalent la présence ou l'absence de symptômes grippaux via cette cohorte en ligne.

Grâce à la flexibilité du système GrippeNet.fr, nous avons pu établir différentes définitions de cas d'épisode grippal à la fois sensibles et spécifiques, afin de comparer les incidences des indicateurs grippaux estimées à partir de GrippeNet.fr à celles issues de deux systèmes de surveillance en soins primaires : le réseau des GROG et le réseau Sentinelles. Les définitions de cas de syndrome grippal, plus spécifiques, étaient comparées aux données du réseau Sentinelles ; celles de l'infection respiratoire aiguë, plus sensibles, aux données du réseau des GROG. Les analyses ont d'ailleurs été menées sur les saisons 2012/2013 et 2013/2014, lorsque les deux réseaux de surveillance sentinelle cohabitaient en France. La différence majeure observée entre les définitions de syndrome grippal et celles d'IRA était liée à l'amplitude du signal obtenu. L'amplitude était plus grande pour les IRA que pour les syndromes grippaux, renvoyant à la sensibilité et spécificité des deux indicateurs respectivement.

En plus de l'importance de l'estimateur grippal à identifier, les critères tenant compte de la déclaration et de la participation des personnes sont primordiaux pour produire des estimations de l'incidence des maladies liées à la grippe. En effet, les participants ne sont pas initiés, l'auto-déclaration faite de leur état de santé n'est pas validé par un professionnel de santé, et il est apparu que les participants soient plus enclins à rapporter des symptômes que l'absence de symptômes.

Alors que les sources d'information sont différentes, l'évolution temporelle des maladies grippales mesurée à partir de GrippeNet.fr était en bonne adéquation avec les tendances issues

des systèmes de soins primaires, anticipant souvent le pic épidémique, et cela malgré une participation irrégulière des individus. Une plus grande précision pour l'IRA que pour le syndrome grippal était néanmoins observée. La définition du syndrome grippal de l'ECDC, utilisée couramment par les autres systèmes participatifs européens, s'est révélée plus efficace pour le suivi de l'IRA que pour le suivi du syndrome grippal avec les données issues de GrippeNet.fr. Concernant l'intensité des épidémies (i.e. amplitude du signal), celle-ci est intrinsèque au système GrippeNet.fr et ne peut donc être clairement comparée aux autres systèmes traditionnels. En effet, elle nécessite des estimations de seuils d'épidémie et d'intensité de l'épidémie sur plusieurs saisons consécutives pour pouvoir comparer une saison par rapport à l'autre. C'est effectivement l'historique des systèmes traditionnels qui permet de faire savoir aux autorités sanitaires si l'épidémie de la saison en cours est d'intensité faible, modérée ou forte, pour que des mesures de contrôle adéquates soient appliquées.

Notre étude a finalement permis de fournir un cadre standardisé à la surveillance participative de la grippe via le système GrippeNet.fr, en montrant sa grande sensibilité à détecter les changements liés aux maladies grippales dans la population. Ce cadre permettra d'aider à la comparaison des différentes saisons, d'autant que huit ont déjà été réalisées, mais aussi à comparer l'épidémie observée en population à celle observée via le système de soins. GrippeNet.fr pourrait être un système complémentaire pour surveiller directement les maladies grippales et plus spécifiquement les IRA à partir de la population générale, d'autant que depuis 2011 l'OMS promeut une surveillance complète de la grippe à différents niveaux [186].

4.3. Article : The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France

The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France

Caroline Guerrisi¹, Clément Turbelin¹, Cécile Souty¹, Chiara Poletto¹, Thierry Blanchon¹, Thomas Hanslik^{1,2,3}, Isabelle Bonmarin⁴, Daniel Levy-Bruhl⁴, Vittoria Colizza¹

1. Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP), Paris, France

2. UFR des sciences de la santé Simone-Veil, Université Versailles-Saint-Quentin-en-Yvelines, Versailles, France

3. AP-HP, Service de Médecine Interne, Hôpital Ambroise Paré, Boulogne Billancourt, France

4. Department of infectious diseases, Public Health France, Saint-Maurice, France

Correspondence: Vittoria Colizza (vittoria.colizza@inserm.fr)

Citation style for this article:

Guerrisi Caroline, Turbelin Clément, Souty Cécile, Poletto Chiara, Blanchon Thierry, Hanslik Thomas, Bonmarin Isabelle, Levy-Bruhl Daniel, Colizza Vittoria. The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France. *Euro Surveill.* 2018;23(25):pii=1700337. <https://doi.org/10.2807/1560-7917.ES.2018.23.25.1700337>

Article submitted on 22 May 2017 / accepted on 15 Feb 2018 / published on 21 Jun 2018

Introduction: Participatory surveillance systems provide rich crowdsourced data, profiling individuals and their health status at a given time. We explored the usefulness of data from GrippeNet.fr, a participatory surveillance system, to estimate influenza-related illness incidence in France. **Methods:** GrippeNet.fr is an online cohort since 2012 averaging ca. 5,000 weekly participants reporting signs/symptoms suggestive of influenza. GrippeNet.fr has flexible criteria to define influenza-related illness. Different case definitions based on reported signs/symptoms and inclusions of criteria accounting for individuals' reporting and participation were used to produce influenza-related illness incidence estimates, which were compared to those from sentinel networks. We focused on the 2012/13 and 2013/14 seasons when two sentinel networks, monitoring influenza-like-illness (ILI) and acute respiratory infections (ARI) existed in France. **Results:** GrippeNet.fr incidence estimates agreed well with official temporal trends, with a higher accuracy for ARI than ILI. The influenza epidemic peak was often anticipated by one week, despite irregular participation of individuals. The European Centre for Disease Prevention and Control ILI definition, commonly used by participatory surveillance in Europe, performed better in tracking ARI than ILI when applied to GrippeNet.fr data. **Conclusion:** Evaluation of the epidemic intensity from crowdsourced data requires epidemic and intensity threshold estimations from several consecutive seasons. The study provides a standardised analytical framework for crowdsourced surveillance showing high sensitivity in detecting influenza-related changes in the population. It contributes to improve the comparability of epidemics across seasons and with sentinel systems. In France, GrippeNet.fr may supplement the

ILI sentinel network after ARI surveillance discontinuation in 2014.

Introduction

Influenza surveillance systems aim to annually track influenza epidemics, detecting their start, monitoring their spatio-temporal spread, identifying populations at risk and circulating viruses, and estimating the impact on the community and healthcare structures [1]. Sentinel surveillance systems are based on primary care and report the weekly number of patients examined with influenza-related illness. Recently, a variety of non-traditional surveillance approaches have emerged [2-9] where large amounts of crowdsourced digital data produced by individuals enable such individuals to contribute to monitoring the health of their community and to provide authorities with additional characterisations of the epidemic. In the European Union, more than one third of the countries run a participatory surveillance system to monitor influenza epidemics, under the standardised framework of the Influenzanet network established in 2011 [2,10,11]. The Influenzanet network is a syndromic surveillance system based on voluntary self-reporting of symptoms by participants residing in countries, which are part of the Influenzanet.

Crowdsourced data bring novel issues regarding data analysis, due to their non-traditional nature. They refer to the dynamic participation of individuals having variable reporting behaviours along the season, individuals' interpretation of the terms used for surveillance, and the correctness of their self-assessments. A small number of these aspects have been analysed in isolation in previous work [12-15], however a systematic evaluation is still missing. Crowdsourced surveillance

TABLE 1

Influenza-like-illness (ILI) and acute respiratory infection (ARI) case definitions used according to the surveillance systems, France, 2012/13 and 2013/14

Case definitions according to the surveillance system ^a		Sudden onset of symptoms	Fever ^b and general symptoms	Respiratory signs
ILI case definition^a				
Sentinel network (RS)		Yes	Fever $\geq 39^{\circ}\text{C}$ AND myalgia ^c	Yes
GrippeNet.fr ^a	ILI	Yes	Fever $\geq 39^{\circ}\text{C}$ AND pain ^c	Sore throat OR cough OR shortness of breath OR sneezing OR rhinorrhoea (runny or blocked nose)
	ILIf	Yes	Fever ^b OR fever $\geq 39^{\circ}\text{C}$ AND pain ^c	Sore throat OR cough OR shortness of breath OR sneezing OR rhinorrhoea (runny or blocked nose)
	ILI-	Yes	Fever ^b OR fever $\geq 38^{\circ}\text{C}$ AND pain ^c OR headache	Sore throat OR cough OR shortness of breath
	ILI-f	Yes	Fever ^b AND pain ^c OR headache	Sore throat OR cough OR shortness of breath
ARI case definition^a				
Regional Influenza Surveillance Group (GROG)		Yes	At least one sign suggesting an acute infection (fever or asthenia or headache or muscle pain, etc.)	Cough OR rhinitis OR coryza
GrippeNet.fr ^a	ARI	Yes	Fever ^b OR chills OR malaise (feeling tired or exhausted) OR headache OR pain ^c	Cough OR rhinorrhoea OR sneezing
	ARI+	Yes	Fever ^b OR chills OR malaise OR headache OR pain ^c	Cough OR rhinorrhoea OR sneezing OR sore throat OR shortness of breath OR coloured sputum/phlegm
	ECDC	Yes	Fever ^b OR chills OR malaise OR headache OR pain ^c	Sore throat OR cough OR shortness of breath

ECDC: European Centre for Disease Prevention and Control.

^a Case definitions and inclusion criteria can vary in the participatory surveillance system GrippeNet.fr, so several possibilities are considered for this system.

^b Participants are free to report fever with or without measure of body temperature. In the table, the mention of 'fever' without a defining temperature threshold refers to self-reported fever without body temperature specification. Fever, without temperature specification, was considered in this study for the GrippeNet.fr case definitions because participatory surveillance systems are adopted in several countries in Europe and depending on individual preferences/habits may not always obtain the measures of body temperature.

^c For individuals above 5 years-old.

is further complicated by the choice of the appropriate definition used to identify influenza cases. Because of the lack of specificity of influenza symptoms, national sentinel systems adopted influenza-like-illness (ILI) or acute respiratory illness (ARI) definitions, two of the most common quantitative indicators, however these indicators are defined at country level and no defined standard exists at the international level [16-18]. The evaluation and comparison of ILI and ARI quantitative indicators are critical for the sensitivity and specificity of resulting public health recommendations, though their performance in estimating the influenza-related illness epidemic characteristics non-trivially depends on age of the cases, circulating influenza subtypes, or medical practices [16,17]. Participatory systems have the advantage of being flexible in the case definitions to adopt, as case definitions can be built on different combinations of the symptoms collected, without requiring an a priori definition to be used by sentinel practitioners. Previous Influenzanet studies have mainly focused on a single case definition [14,15], thus it remains unclear how different combinations of self-reported symptoms perform in the accurate estimation

of influenza incidences, which can be compared to available sentinel estimates.

Here we considered the case of France, where the participatory system GrippeNet.fr (GN) was established in 2012 as part of Influenzanet [19]. Previous work focused on individual-level epidemiological analyses, generally not possible in sentinel networks, allowed by the availability of individual data on demographic indicators, lifestyle, health variables, and attitudes [10,20,21]. In this study we focused on population-level indicators, and analysed crowdsourced incidence estimates comparing them to official estimates provided by sentinel systems. The 2012/13 and 2013/14 seasons were chosen for this work, because during this period two primary care surveillance systems monitoring independently ILI and ARI were available, against which we could assess the accuracy of GN analyses.

We proposed and assessed different inclusion criteria accounting for individuals' reporting and participation, and evaluated the accuracy of a set of case definitions. In addition, in light of the termination of the ARI sentinel

TABLE 2

GrippeNet.fr (GN) incidence computation criteria, France

Computation criteria		Aim	Description
Raw criteria		Basic computation on raw data	All ILI/ARI episodes are considered. Weekly incidence rate is computed as the number of ILI/ARI episodes in the week divided by the total number of GN participants in the cohort from the beginning of the season up to that week
Episode reporting criteria	First survey exclusion	To account for first-time reporting bias	Weekly incidence rate is computed in the same way as for the raw criteria, but with the exclusion of the first report of each new participant
	Episode merging	To exclude reports on clinically unlikely different episodes	As raw, with the merging of consecutive ILI/ARI episodes within 2 weeks of each other
Participation criteria	Minimum number of symptoms reports (<i>m</i>)	To exclude individuals who have reported rarely	Denominator in the incidence computation is equal to the number of participants who filled $\geq m$ reports, $m = 2, 3$
	Participation window	To discard irregular participation and define a time window during which a participant can be considered actively engaged in the study	Denominator in the incidence computation is equal to the number of participants who reported at least once in the time window of <i>n</i> weeks before and after the reporting week; $n = 0, 1, 2, 3, 4$ ($n = 0$ means that participants are counted only in the week of reporting; $n = 1$ means that participants are counted in the week of reporting and also in the week before and in the week after)

ARI: acute respiratory infection; ILI: influenza-like-illness.

surveillance system since 2014, we evaluated the possibility to use GN as a replacement for ARI surveillance. The overall aim was to: interpret GN weekly incidence estimates for influenza-related illness compared with available sentinel knowledge to provide continuous and robust data to monitor epidemic trends and intensities; propose a standardised approach to compute crowdsourced incidence, which could be extended to other Influenzanet countries allowing comparability across them and with national systems; learn from the only two seasons in France where three surveillance systems were in place (2 sentinel systems and 1 participative system) to assess GN’s potential benefits as an adjunct to the ILI sentinel influenza network in France.

Methods

GrippeNet.fr data collection

GN is a participatory surveillance system collecting voluntary reports of influenza-related symptoms through a dedicated website (<https://www.grippenet.fr>) where individuals also provide profile information [19]. Data are collected on a weekly basis through a symptoms survey [2,10,11]. Individuals are asked if they experienced any of the following 19 symptoms (or they can opt to report ‘no symptoms’): fever, chills, runny or blocked nose, sneezing, sore throat, cough, shortness of breath, headache, muscle/joint pain, chest pain, feeling tired or exhausted, loss of appetite, coloured sputum, watery/bloodshot eyes, nausea, vomiting, diarrhoea, stomach ache, other symptoms. If symptoms are reported, further questions are asked to characterise the participant’s behaviour (general practitioner (GP) consultation, drugs uptake) and to assess the syndrome (e.g. sudden onset, onset date).

For reports of fever, a further question concerning the level of fever is asked, where the participant can input a measure of body temperature, however the response to this question is optional, so the temperature for fever is not always obtained. The full survey is provided as supplementary material (Supplement).

With respect to the general population, participants in GN are on average older and include a larger proportion of women [19]. GN population is however representative in terms of health indicators such as diabetes and asthma conditions.

French surveillance networks

Two sentinel surveillance systems have been operating in France to monitor the influenza circulation in the country: Réseau Sentinelles (sentinel network, RS) estimating the weekly number of ILI cases [22]; GROG (Regional Influenza Surveillance Group) estimating the weekly number of ARI cases [23]. In 2014 GROG was discontinued and only ILI is currently monitored in France. Case definitions are reported in Table 1.

Drug sales data

We considered sales data for 14 classes of medications in the EphMRA classification of interest for influenza [24]. Drug sales data are based on a representative sample of 14,000 pharmacies accounting for 60% of pharmacies in France.

GrippeNet.fr data analyses: case definitions

Different case definitions can be built from GN symptoms (Table 1). We considered the GN_{ILI} and GN_{ARI} definitions most closely matching ILI and ARI definitions adopted by RS and GROG, respectively. We explored variations of GN_{ILI} by: relaxing the constraint

TABLE 3

GrippeNet.fr seasons' descriptive results, France, 2012/13 and 2013/14

Parameters	Criteria	2012/13	2013/14
		Week 46 2012–week 16 2013 (23 weeks)	Week 46 2013–week 15 2014 (22 weeks)
Number of participants	- Raw	6,046	5,907
Number of total symptoms survey	- Raw	77,875	83,455
Mean number of participants per week (± standard deviation)	- Raw	4,987 (±1,224)	5,112 (±725)
	- First survey excluded ^a +episodes merged ^b + $m=2$	4,192 (±1,039)	4,634 (±665)
	- First survey excluded ^a +episodes merged ^b + $m=3$	3,942 (±914)	4,466 (±605)
	- First survey excluded ^a +episodes merged ^b + $m=2+n=0$	2,831 (±574)	3,326 (±353)
	- First survey excluded ^a +episodes merged ^b + $m=2+n=1$	3,568 (±718)	4,127 (±452)
	- First survey excluded ^a +episodes merged ^b + $m=2+n=2$	3,748 (±775)	4,297 (±489)
	- First survey excluded ^a +episodes merged ^b + $m=2+n=3$	3,839 (±814)	4,375 (±512)
	- First survey excluded ^a +episodes merged ^b + $m=3+n=0$	2,813 (±575)	3,312 (±357)
	- First survey excluded ^a +episodes merged ^b + $m=3+n=1$	3,520 (±709)	4,090 (±455)
	- First survey excluded ^a +episodes merged ^b + $m=3+n=2$	3,672 (±756)	4,240 (±487)
	- First survey excluded ^a +episodes merged ^b + $m=3+n=3$	3,740 (±785)	4,302 (±506)

m: minimum number of reports of symptoms; *n*: number of participation weeks around the reporting week.

^a First report of each new participant is excluded.

^b Influenza-like illness/acute respiratory infections (ILI/ARI) episodes experienced within 2 weeks of previous ILI/ARI episodes are considered to belong to the same illness episode.

on body temperature, whereby participants who declared a fever episode without specifying their body temperature were included in addition to participants reporting a body temperature $\geq 39^\circ\text{C}$ (GN_{ILI^+}); considering a lower temperature cut-off ($\geq 38^\circ\text{C}$), the inclusion of ‘headache’, and a restricted set of respiratory symptoms (GN_{ILI^-}); removing the constraint on body temperature (GN_{ILI^-}). For the GN_{ARI} definition, we considered: a larger set of respiratory symptoms (GN_{ARI^+}) as well as the case definition of the European Centre for Disease Control and Prevention (GN_{ECDC}), which is often used by Influenzanet studies [14,15]. Though formally an ILI definition, we included GN_{ECDC} in the ARI classification as it is less restrictive in the inclusion of ‘fever’ [25].

GrippeNet.fr data analyses: incidence computation

We focused on the 2012/13 (from week 46 2012 to week 16 2013) and 2013/14 (from week 46 2013 to week 15 2014) influenza seasons. For each case definition and season, we estimated weekly incidence rates of cases from GN raw data considering the combination of several criteria (Table 2). The raw incidence time series is computed as the ratio between the number of

GN participants declaring an episode in a given week and the total number of participants registered in the cohort from the beginning of the season until the week considered.

Subsequently, we first considered criteria accounting for different behaviours in reporting illness episodes. The *first survey exclusion* adjusts GN incidence rates by removing the result of the first survey of newly enrolled participants [5,13-15,26], as participants are more prone to report symptoms at their first report following enrolment [13,14]. To account for the lack of validation by a GP, the *episodes merging* criterion considers ILI/ARI episodes experienced within 2 weeks of previous ILI/ARI episodes to be part of the same illness episode [12,26]. Second, we considered inclusion criteria accounting for heterogeneous participation. We implemented a *minimum number *m* of symptoms reports* per individual throughout the season ($m=2$ reports or $m=3$ reports, including the first survey) to discard those with rare participation [13]. We considered the inclusion criterion of a *participation window* of *n* weeks around the reporting week ($n=0, 1, 2, 3, 4$) to account for non-continuous participation. If $n=0$ each participant

TABLE 4

Performance of computation methods in the comparison between GrippeNet.fr influenza-like illness and acute respiratory infection incidence estimates and official sentinel network estimates, France, 2012/13 and 2013/14

Criteria for participation and behaviour considered when analysing GrippeNet.fr data		2012/13		2013/14	
		Influenza-like illness	Acute respiratory infection	Influenza-like illness	Acute respiratory infection
Raw	<i>r</i>	0.57	0.69	0.81	0.86
	<i>e</i>	3.4	0.8	3.7	0.4
First survey excluded ^a	<i>r</i>	0.87	0.81	0.82	0.85
	<i>e</i>	1.0	0.2	2.2	0.2
First survey excluded ^a + episodes merged ^b	<i>r</i>	0.84	0.81	0.80	0.84
	<i>e</i>	0.9	0.2	1.8	0.1
First survey excluded ^a + episodes merged ^b + $m=2$	<i>r</i>	0.89	0.89	0.83	0.89
	<i>e</i>	1.0	0.2	2.0	0.1
First survey excluded ^a + episodes merged ^b + $m=3$	<i>r</i>	0.88	0.90	0.82	0.89
	<i>e</i>	1.0	0.2	2.0	0.1
First survey excluded ^a + episodes merged ^b + $m=2+n=0$	<i>r</i>	0.90	0.92	0.85	0.95
	<i>e</i>	1.7	0.7	3.4	0.7
First survey excluded ^a + episodes merged ^b + $m=2+n=1$	<i>r</i>	0.90	0.92	0.85	0.93
	<i>e</i>	1.2	0.3	2.4	0.3
First survey excluded ^a + episodes merged ^b + $m=2+n=2$	<i>r</i>	0.90	0.92	0.85	0.92
	<i>e</i>	1.1	0.2	2.2	0.2
First survey excluded ^a + episodes merged ^b + $m=2+n=3$	<i>r</i>	0.90	0.92	0.85	0.91
	<i>e</i>	1.0	0.2	1.4	0.2
First survey excluded ^a + episodes merged ^b + $m=3+n=0$	<i>r</i>	0.88	0.90	0.84	0.94
	<i>e</i>	1.7	0.7	3.2	0.6
First survey excluded ^a + episodes merged ^b + $m=3+n=1$	<i>r</i>	0.89	0.91	0.84	0.92
	<i>e</i>	1.2	0.3	2.3	0.2
First survey excluded ^a + episodes merged ^b + $m=3+n=2$	<i>r</i>	0.89	0.91	0.84	0.91
	<i>e</i>	1.1	0.2	2.1	0.2
First survey excluded ^a + episodes merged ^b + $m=3+n=3$	<i>r</i>	0.88	0.91	0.84	0.91
	<i>e</i>	1.1	0.2	2.1	0.2

ARI: acute respiratory infection; *e*: mean error; ILI: influenza-like illness; *m*: minimum number of reports of symptoms; *n*: number of participation weeks around the reporting week; *r*: mean correlation.

In the table, incidence estimates based on the various ILI or ARI definitions from GrippeNet.fr were each used to compute respective correlations and errors against incidence estimates based on sentinel ILI and sentinel ARI. For each criteria for participation and behaviour in the table, a mean correlation and error across all GrippeNet.fr ILI definitions, and a mean correlation and error across all GrippeNet.fr ARI definitions are respectively presented.

^a First report of each new participant is excluded.

^b ILI/ARI episodes experienced within 2 weeks of previous ILI/ARI episodes are considered to belong to the same illness episode.

is counted only in the week of reporting; if $n=1$ each participant is considered to be part of the cohort in the week of reporting and also in the week before and the week after that, assuming e.g. that they forgot to connect online (analogously for $n>1$). These criteria were combined in a stepwise progression (see results).

Incidence time series computed on these datasets were adjusted by age group (0–14, 15–44, 45–64, ≥65 years) to account for the non-representative nature of the GN population [19], and smoothed through a linear filtering method [27] commonly used to filter out undesired spikes induced by large variations in enrolment (e.g. following a communication action). They were computed at the national and regional level (Ile-de-France,

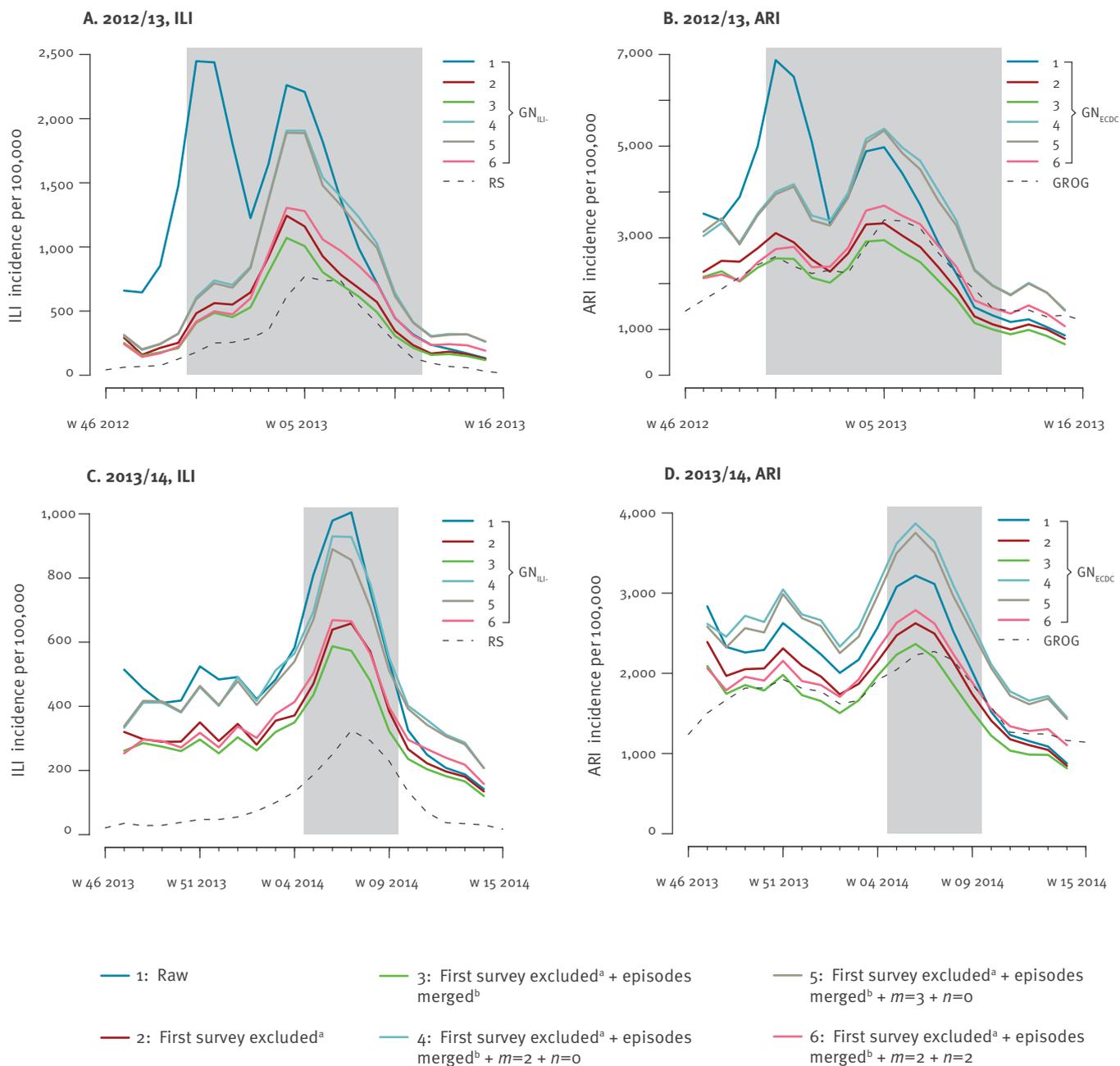
North-East, North-West, South-East and South-West, see Figure S1 in Supplement) with standard adjustment by region.

Statistical analysis

GN incidence time series were compared with the official estimates of RS and GROG through the Pearson correlation coefficient (*r*) using the Bonferroni correction ($p<0.01$) to evaluate the agreement of time series trends, and the normalised root mean square error (*e*, error divided by the average to compare measurements with different scales) to evaluate the agreement of influenza intensity. GN incidence trends were also compared with the drug sales time series through the correlation *r*.

FIGURE 1

Comparison between official time series and GrippNet.fr incidence curves using different computation methods, France 2012/13 and 2013/14



ARI: acute respiratory infection; GN: GrippNet.fr; GROG: regional influenza surveillance group estimating the weekly number of ARI cases; ILI: influenza-like illness; RS: sentinel network estimating the weekly number of ILI cases; m: minimum number of reports of symptoms; n: number of participation weeks around the reporting week; w: week.

A to D show different computation methods for 2012/13 season (A, B) and 2013/14 season (C, D). A and C show GrippNet.fr ILI (GNILI-, coloured continuous lines) and RS (dashed black lines); B and D show GrippNet.fr ARI (GNECDC, coloured continuous lines) and GROG (dashed black lines). Shaded areas in all panels represent the 2012/13 and 2013/14 epidemic periods reported by RS.

^a First report of each new participant is excluded.

^b ILI/ARI episodes experienced within 2 weeks of previous ILI/ARI episodes are considered to belong to the same illness episode.

TABLE 5

Performance of case definitions in the comparison between GrippeNet.fr incidence estimates and official estimates, France, 2012/13 and 2013/14

Comparison	ILI and ARI case definitions	2012/13			2013/14		
		Correlation r (95% CI)	Error e	Peak week difference ^a	Correlation r (95% CI)	Error e	Peak week difference ^a
GN ILI	GN _{ILI} vs RS	0.89 (0.75–0.96)	0.62	2	0.84 (0.64–0.94)	0.51	0
	GN _{ILIF} vs RS	0.81 (0.58–0.92)	0.56	-1	0.83 (0.61–0.93)	1.40	-1
vs sentinel ILI	GN _{ILI-} vs RS	0.95 (0.87–0.98)	1.04	-1	0.91 (0.79–0.97)	2.35	-1
	GN _{ILI-f} vs RS	0.94 (0.86–0.98)	2.06	0	0.83 (0.62–0.93)	4.58	-1
GN ARI	GN _{ARI} vs GROG	0.91 (0.79–0.96)	0.20	-1	0.93 (0.82–0.97)	0.18	-1
	GN _{ARI+} vs GROG	0.90 (0.78–0.96)	0.30	0	0.90 (0.76–0.96)	0.31	-1
vs sentinel ARI	GN _{ECDC} vs GROG	0.96 (0.89–0.98)	0.13	0	0.93 (0.83–0.97)	0.16	-1
GN ILI	GN _{ILI} vs GROG	0.77 (0.51–0.90)	0.95	2	0.84 (0.64–0.94)	0.94	0
	GN _{ILIF} vs GROG	0.82 (0.61–0.93)	0.87	-1	0.86 (0.68–0.94)	0.87	-1
vs sentinel ARI	GN _{ILI-} vs GROG	0.81 (0.59–0.92)	0.77	-1	0.86 (0.68–0.95)	0.80	-1
	GN _{ILI-f} vs GROG	0.88 (0.73–0.95)	0.63	0	0.89 (0.73–0.95)	0.67	-1
GN ARI	GN _{ARI} vs RS	0.76 (0.49–0.90)	7.9	-1	0.66 (0.31–0.85)	17.5	-1
	GN _{ARI+} vs RS	0.76 (0.49–0.90)	8.7	0	0.60 (0.21–0.82)	19.7	-1
vs sentinel ILI	GN _{ECDC} vs RS	0.85 (0.66–0.94)	7.4	0	0.66 (0.31–0.85)	17.0	-1

ARI: acute respiratory infection; CI: confidence interval; GN: GrippeNet.fr; GROG: regional influenza surveillance group estimating the weekly number of ARI cases; ILI: influenza-like illness; RS: sentinel network estimating the weekly number of ILI cases.

The different case definitions used in GN (GN_{ILI}, GN_{ILIF}, GN_{ILI-}, GN_{ILI-f}, GN_{ARI}, GN_{ARI+}, GN_{ECDC}) are described in Table 1.

^a The difference is calculated as peak week of GN incidence minus peak week of official estimate.

Ethics statement

GN was reviewed and approved by the French Advisory Committee for research on information treatment in the field of health (i.e. CCTIRS, authorisation 11.565), and by the French National Commission on Informatics and Liberty (i.e. CNIL, authorisation DR-2012-024).

Results

A total of 6,046 individuals participated to GN submitting 77,875 symptom reports during the 2012/13 season (raw data, Table 3); 65% of them also participated in the 2013/14 season, which reached a total of 5,907 participants with 83,455 surveys filled in. The system had a mean weekly participation of 4,987 ($\pm 1,224$, standard deviation) and 5,112 (± 725) individuals in the two seasons. Participation criteria led to a progressive reduction of these numbers, stronger for the first season under study (25% reduction in 2012/13 vs 16% in 2013/14 with first survey exclusion, episodes merging, and $m = 2 + n = 2$ method compared with raw data).

GrippeNet.fr incidence computation methods

Overall, across all definitions of IRI and ARI respectively used based on GrippeNet.fr data, correlation was moderate ($r = 0.57 \pm 0.29$) and strong ($r = 0.69 \pm 0.01$) in the 2012/13 season for ILI and ARI, respectively, and very strong for both indicators in the 2013/14 season

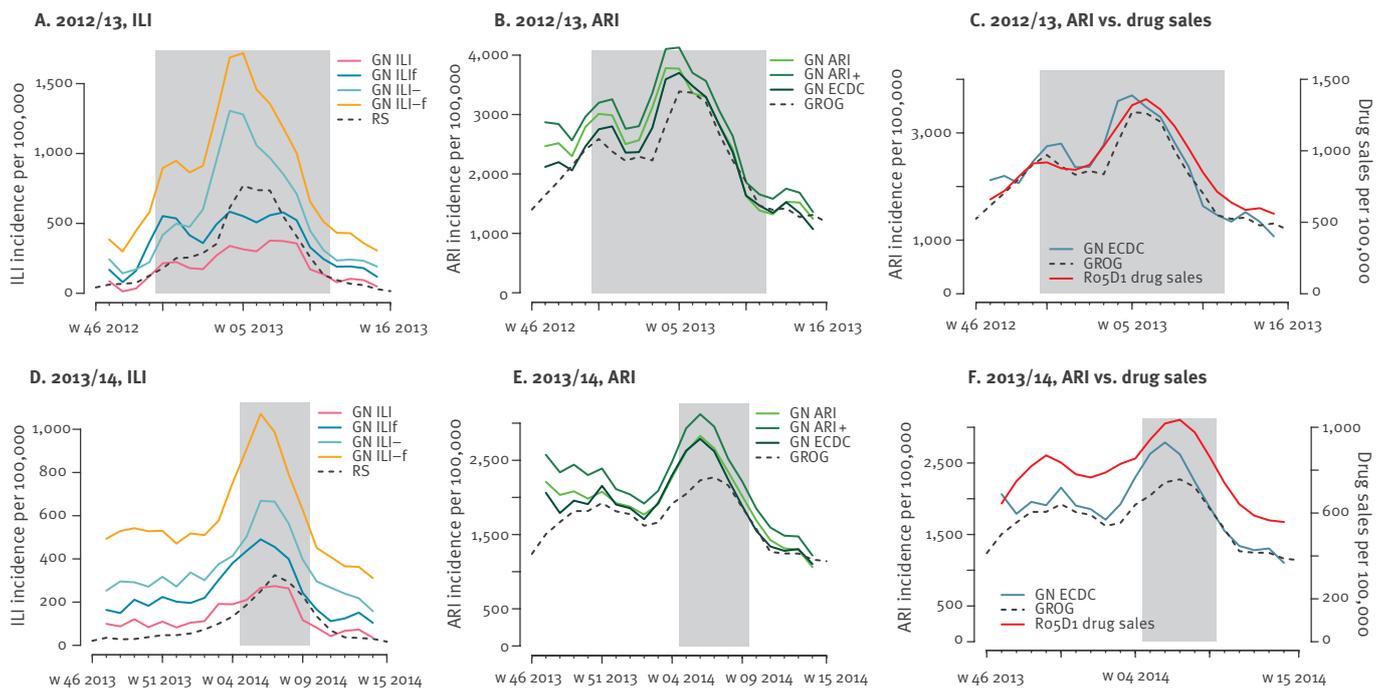
($r = 0.81 \pm 0.03$ for ILI, $r = 0.86 \pm 0.02$ for ARI), when calculated on raw data (Table 4). Errors were much larger for ILI than for ARI comparisons ($e = 3.4 \pm 1.69$ for ILI vs $e = 0.8 \pm 0.06$ for ARI in 2012/13, and similarly for 2013/14 season).

Excluding the first report of newly enrolled participants strongly increased correlation in the first season under study, whereas no effect was reported in the second season. Errors were considerably reduced in both seasons and for both indicators. This improvement is obtained as the *first survey exclusion* removes the large epidemic peak observed at the beginning of 2012/13 season that is not reported by sentinel sources, as illustrated in Figure 1 (panels A,B) for the particular examples of incidence estimates based on the GN_{ILI-} and GN_{ECDC} ILI definitions. Merging close consecutive episodes did not visibly affect incidence estimates. Adding inclusion criteria on participation of individuals slightly improved correlations, however strongly increasing the errors in the case $n = 0$ (Table 4), as it reduced the number of participants in the cohort (Table 3).

Comparison between official time series and GrippeNet.fr incidence curves using different computation methods, France 2012/13 and 2013/14

FIGURE 2

Comparison between official time series and Grippenet.fr incidence curves using different case definitions within Grippenet.fr, France 2012/13 and 2013/14



ARI: acute respiratory infection; ECDC: European Centre for Disease Prevention and Control; GN: Grippenet.fr; GROG: regional influenza surveillance group estimating the weekly number of ARI cases; ILI: influenza-like illness; RS: sentinel network estimating the weekly number of ILI cases; w: week.

The different case definitions used in GN (GN_{ILI}, GN_{ILIF}, GN_{ILI-}, GN_{ILI-f}, GN_{ARI}, GN_{ARI+}, GN_{ECDC}) are described in Table 1.

A, B, D, E show different case definitions for the 2012/13 season (A, B) and the 2013/14 season (D, E). A and D show Grippenet.fr ILI case definitions (coloured continuous lines) and RS (dashed black line); B and E show Grippenet.fr ARI case definitions (coloured continuous lines) and GROG (dashed black line). C and F show GN_{ECDC} case definition, GROG, and drug sales. Shaded areas in all panels represent the 2012/13 and 2013/14 epidemic periods reported by RS.

Case definitions

To assess the accuracy of case definitions, we considered a baseline computation method composed of the episodes reporting criteria and participation criteria with $m=2$ and $n=2$. The largest correlation was obtained in both seasons with GN_{ILI-} for ILI indicator with $r=0.95$ (95% confidence interval (CI): 0.87–0.98) in 2012/13, and $r=0.91$ (95%CI: 0.79–0.97) in 2013/14, Table 5) and with GN_{ECDC} for ARI indicator ($r=0.96$; 95%CI: 0.89–0.98 in 2012/13, $r=0.93$ (95%CI: 0.83–0.97) in 2013/14). Both case definitions brought an anticipation of the peak time of one week, with the exception of GN_{ECDC} in 2012/13 season. The overall smallest correlation was observed for GN_{ILIF} vs RS in both seasons.

A higher agreement in the incidence trend and smaller errors were reported for ARI case definitions (e in the range 0.13–0.31) compared with ILI (0.51–4.58), with ARI curves rather close to each other and to GROG time series (Figure 2).

The cross-comparison between ILI indicators from GN and ARI estimates provided by GROG led to strong correlations, slightly lower than those obtained between ILI indicators from GN and RS estimates (Table 5). ARI estimates from GN instead generally performed more poorly when compared with RS estimates ($r=0.85$ vs 0.96 for GN_{ECDC} vs RS compared with GN_{ECDC} vs GROG in 2012/13; $r=0.66$ vs 0.93 in 2013/14).

The GN_{ILI-} and GN_{ECDC} case definitions for ILI and ARI, respectively, yielded the largest correlation with the plain antitussive drug class (Ro5D1, $0.89 \leq r \leq 0.92$). The Ro5D1 sales trend reports a first peak early in the season, similarly to what estimated by GN (Figure 2; C, F).

Regional incidence estimates

Correlations between GN regional estimates and corresponding official estimates are generally lower than the ones obtained at the national level, ranging from 0.61 in the North-West region to 0.87 in South-East in 2012/13, and from 0.64 in North-West, North-East, South-East to 0.75 in Ile-de-France in 2013/14 (Table

S1 in Supplement). Weak non-significant associations were found between correlation coefficients and regional participations.

Discussion

Despite the intrinsic differences between the self-selected population of GN and the medically assisted population considered by sentinel surveillances, epidemic trends from GN reports compare well with official sources. Notably, varying participation criteria led to uniformly very high correlations, suggesting that once adjusted for first-time reporting bias, crowdsourced indicators accurately summarise the seasonal epidemics, regardless of the regularity of participation of its volunteers.

Epidemic intensities, on the other hand, are strongly dependent on participation criteria, as these alter the time-dependent size of the cohort. Higher intensities are expected in GN estimates compared with sentinel surveillance as the system also includes individuals not seeking medical care [20]. An appropriate frame of reference is currently missing because of the rather recent availability of this type of surveillance data, hindering the comparability of epidemic intensity across seasons and across surveillance systems. This still remains an unresolved challenge also in the framework of traditional sentinel networks in Europe, with countries reporting intensity levels based on national references, obtained with different methods, and depending on patterns of medical care consultation [28].

Given an arbitrary set of inclusion criteria (e.g. $m=2$, $n=2$) proposed here as a new standard to analyse crowdsourced data, an automated standardised method to define the epidemic threshold and a set of intensity thresholds should be put in place for the comprehensive analysis of crowdsourced data. For France, such implementation will soon be possible, after accumulating a long enough historical dataset from crowdsourced surveillance for thresholds estimation (minimum of six consecutive seasons for the moving epidemic method adopted in [28]). The advantage compared with traditional sources would be the rather seamless introduction of the method across all Influenzanet countries without additional burden on primary care surveillance, thus allowing the understanding of epidemic patterns across seasons and countries independently of health-seeking behaviour.

All case definitions performed rather well ($r>0.81$) when comparing ILI case definitions on crowdsourced data with ILI sentinel surveillance, and analogously for ARI case definitions. Higher correlations were consistently obtained when less restrictive case definitions were considered. For ILI, lower temperature cutoffs or the addition of 'headache' in the general symptoms considerably improved the accuracy of GN case definition, even though headache is not a symptom which is associated with ILI by GPs in France. Previous work identified body temperature to be independently

associated with virologically confirmed influenza with an increasing likelihood associated to rising temperature [16]. We found self-reported fever to be required as a mandatory criterion to accurately monitor ILI, in line with previous evidence [16,17], with best results obtained for body temperatures $\geq 38^{\circ}\text{C}$. All case definitions having fever as an optional criterion better reproduce ARI incidence trends, confirming previous results on influenza cases [16,17]. Interestingly, comparisons with estimates from GROG, using a less restrictive case definition, were generally more accurate than with RS, adopting a more specific definition.

All GN estimates reported the small peak of cases likely corresponding to the circulation of respiratory syncytial virus (RSV) viruses in the early weeks of the influenza epidemic period, visible only in ARI sentinel activity and identifiable by increased sales of antitussive drug class generally used for treatment. All these results seem to indicate that participants may find it easier to identify systemic symptoms rather than specific ones, given their lack of medical background. Interpretation of terms used in the case definitions may indeed be different between sentinel practitioners and the general population, and this may considerably impact the measure of influenza activity, as discussed in [18].

Our study shows that the ECDC definition, though formally an ILI definition, is a reliable definition for ARI but it performs rather poorly in tracking ILI in the general population. This finding is essential for participatory surveillance as the ECDC case definition is the one in use in Influenzanet. While a common case definition across countries is critical to standardise comparisons, different needs may emerge at the national level to monitor alternative indicators. Aguilera et al. [18] proposed sentinel networks to introduce a standardised definition to ensure comparability and compatibility of data, in addition to the domestic one historically adopted by each country. Based on self-reported symptoms that are digitally collected in real-time, participatory systems may easily introduce a double case definition, one for domestic use plus a standardised one, overcoming the challenges of increased workload of primary care surveillance.

The accuracy of GN in tracking ARI cases in the general population is specifically important in France, given that the ARI sentinel network was discontinued in 2014 due to changes in influenza surveillance practice. In this context, participatory surveillance may offer an alternative approach to continue the syndromic surveillance of ARI in the country, ensuring a high sensitivity in detecting changes in the epidemic throughout the season. In addition, GN could be coupled in the near future with virological confirmation from self-collected swabs from volunteering participants. With a pilot study within the United Kingdom system of Influenzanet [29], recent works confirmed indeed the feasibility and validity of self-collected swabs for respiratory virus surveillance [30,31]. The participatory approach would thus offer a

novel solution to meet the recent recommendation by the World Health Organization prompting national sentinel networks to monitor also ARI episodes in order to describe a broader range of non-influenza viral pathogens [32]. This is particularly relevant in Europe where the large majority of countries (23/29) monitor exclusively ILI [28].

The flexibility of GN in establishing a case definition based on reported symptoms, without requiring the need for an a priori definition for GPs, is a critical difference with sentinel networks. Such flexibility could contribute to the identification and implementation of optimal case definitions per age class [16]. Most importantly, it would allow the adaptation of surveillance in real-time to monitor an unexpected or atypical clinical manifestation in the context of emerging influenza viruses. The system would also be able to offer continuous surveillance during a pandemic emergency when the public health infrastructure is expected to be overburdened.

An anticipation of 1 week in the incidence peak of GN vs sentinel network is found for the ILI case definition displaying the highest correlation, GN_{ILI} , as well as for the majority of case definitions explored. This suggests that self-reported digital data collected directly from the general population and analysed in a timely fashion can considerably reduce the time needed to produce sentinel estimates (that needs to account for the delay for consulting a primary care doctor and centralising sentinel data for analysis). Confirmed by other participatory systems [11,13-15,33], this feature may function as an important alert system for public health preparedness before reaching the highest weekly incidence value.

The study presents some limitations. GN population is not representative of the general population. While we were able to adjust on age and regional geographic distribution, further adjusting along additional indicators would strongly reduce the sample size in the stratification of the population, thus preventing significant analyses. The agreement found with GP incidence trends suggests however that these limitations have little effect once results are adjusted on basic indicators.

The small number of participants at the regional level led to noisier GN estimates. Correlations with official sources were found to be weakly associated with GN participation in the region, though non-significant. It would be interesting to increase the sample size and systematically assess such dependency on the European scale to identify the minimum number of participants per geographical area to obtain reliable results.

The study was conducted on two seasons only as this is the only time period during which two independent sentinel networks existed in France to monitor different

indicators of influenza-related illness against which GN surveillance data could be evaluated.

Conclusion

GN is an online participatory system providing flexibility and a richness of data offering several opportunities to track and analyse epidemics due to pathogens causing influenza-suggestive symptoms. We evaluated the accuracy of seven influenza case definitions and of different inclusion criteria accounting for variable reporting and participating behaviours of individuals in the online cohort. GN estimates are in very good agreement with the official trends, proving the system to be rather sensitive in detecting influenza-related changes in the population and often anticipating the peak, regardless of the regularity of participation. Evaluation of influenza intensity will benefit from the estimation of epidemic and intensity thresholds for crowdsourced surveillance data once enough seasons will become available. Less restrictive case definitions could be chosen for surveillance purposes as they are found to be more accurate. The ECDC definition for ILI currently adopted by the European participatory system Influenzanet would perform poorly in tracking ILI epidemics, therefore we suggest Influenzanet countries to adopt additionally a case definition including fever and body temperature with a low cutoff ($\geq 38^{\circ}\text{C}$) as compulsory criteria. The standardised method we proposed is essential for the comparability and compatibility of crowdsourced estimates allowing the understanding of epidemic patterns across seasons and countries. In France, GN would represent the ideal candidate to supplement the current sentinel network by monitoring ARI after the termination of GROG and prospectively including self-swabbing.

Acknowledgements

We thank all current and former participants and members of the GrippeNet.fr system.

This work is partly supported by the UMR-S 1136/Public Health France partnership and by the HARMSFLU project (contract ANR-12-MONU-0018).

Conflict of interest

None declared.

Authors' contributions

CG, CT, VC designed the study and performed the statistical analyses. CG, CT, CS, CP, TB, TH, IB, DLB, VC interpreted the results. CG, VC wrote the manuscript. CG, CT, CS, CP, TB, TH, IB, DLB, VC reviewed the manuscript.

References

1. World Health Organization (WHO). WHO global technical consultation: global standards and tools for influenza surveillance, 8-10 March 2011. Geneva: WHO; 2011. [Accessed 3 May 2017]. Available from: http://www.who.int/influenza/resources/documents/technical_consultation/en/

2. Influenzaneet. 2013. [Accessed 6 Sep 2016]. Available from: <http://www.influenzaneet.eu/>
3. Flu Tracking. 2007. [Accessed 6 Sep 2016]. Available from: <http://www.flutracking.net/>
4. Reporta. 2009. [Accessed 6 Sep 2016]. Available from: <http://reporta.c3.org.mx/>
5. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons. *Am J Public Health*. 2015;105(10):2124-30. <https://doi.org/10.2105/AJPH.2015.302696> PMID: 26270299
6. Salud Boricua. 2012. [Accessed 6 Sep 2016]. Available from: <https://saludboricua.org/>
7. Dugas AF, Jalalpour M, Gel Y, Levin S, Torcaso F, Igusa T, et al. Influenza forecasting with Google Flu Trends. *PLoS One*. 2013;8(2):e56176. <https://doi.org/10.1371/journal.pone.0056176> PMID: 23457520
8. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. *PLOS Comput Biol*. 2012;8(7):e1002616. <https://doi.org/10.1371/journal.pcbi.1002616> PMID: 22844241
9. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. *Emerg Themes Epidemiol*. 2014;11(1):7. <https://doi.org/10.1186/1742-7622-11-7> PMID: 24991229
10. Guerrisi C, Turbelin C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, et al. Participatory syndromic surveillance of influenza in Europe. *J Infect Dis*. 2016;214(suppl_4):S386-92. <https://doi.org/10.1093/infdis/jiw280> PMID: 28830105
11. Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, et al. Web-based participatory surveillance of infectious diseases: the Influenzaneet participatory surveillance experience. *Clin Microbiol Infect*. 2014;20(1):17-21. <https://doi.org/10.1111/1469-0691.12477> PMID: 24350723
12. Friesema IHM, Koppeschaar CE, Donker GA, Dijkstra F, van Noort SP, Smallegenburg R, et al. Internet-based monitoring of influenza-like illness in the general population: experience of five influenza seasons in The Netherlands. *Vaccine*. 2009;27(45):6353-7. <https://doi.org/10.1016/j.vaccine.2009.05.042> PMID: 19840672
13. Vandendijck Y, Faes C, Hens N. Eight years of the Great Influenza Survey to monitor influenza-like illness in Flanders. *PLoS One*. 2013;8(5):e64156. <https://doi.org/10.1371/journal.pone.0064156> PMID: 23691162
14. Kjelsø C, Galle M, Bang H, Ethelberg S, Krause TG. Influmeter - an online tool for self-reporting of influenza-like illness in Denmark. *Infect Dis (Lond)*. 2016;48(4):322-7. <https://doi.org/10.3109/23744235.2015.1122224> PMID: 26654752
15. Adler AJ, Eames KT, Funk S, Edmunds WJ. Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey. *BMC Infect Dis*. 2014;14(1):232. <https://doi.org/10.1186/1471-2334-14-232> PMID: 24885043
16. Casalegno JS, Eibach D, Valette M, Enouf V, Daviaud I, Behillil S, et al. Performance of influenza case definitions for influenza community surveillance: based on the French influenza surveillance network GROG, 2009-2014. *Euro Surveill*. 2017;22(14):30504. <https://doi.org/10.2807/1560-7917.ES.2017.22.14.30504> PMID: 28422004
17. Jiang L, Lee VJ, Lim WY, Chen MI, Chen Y, Tan L, et al. Performance of case definitions for influenza surveillance. *Euro Surveill*. 2015;20(22):21145. <https://doi.org/10.2807/1560-7917.ES2015.20.22.21145> PMID: 26062645
18. Aguilera JF, Paget WJ, Mosnier A, Heijnen ML, Uphoff H, van der Velden J, et al. Heterogeneous case definitions used for the surveillance of influenza in Europe. *Eur J Epidemiol*. 2003;18(8):751-4. <https://doi.org/10.1023/A:1025337616327> PMID: 12974549
19. Debin M, Turbelin C, Blanchon T, Bonmarin I, Falchi A, Hanslik T, et al. Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France. *PLoS One*. 2013;8(9):e73675. <https://doi.org/10.1371/journal.pone.0073675> PMID: 24040020
20. Ariza M, Guerrisi C, Souty C, Rossignol L, Turbelin C, Hanslik T, et al. Healthcare-seeking behaviour in case of influenza-like illness in the French general population and factors associated with a GP consultation: an observational prospective study. *BJGP Open*. 2017;1(4).
21. Boiron K, Sarazin M, Debin M, Raude J, Rossignol L, Guerrisi C, et al. Opinion about seasonal influenza vaccination among the general population 3 years after the A(H1N1)pdm2009 influenza pandemic. *Vaccine*. 2015;33(48):6849-54. <https://doi.org/10.1016/j.vaccine.2015.08.067> PMID: 26322844
22. Flahault A, Blanchon T, Dorléans Y, Toubiana L, Vibert JF, Valleron AJ. Virtual surveillance of communicable diseases: a 20-year experience in France. *Stat Methods Med Res*. 2006;15(5):413-21. <https://doi.org/10.1177/0962280206071639> PMID: 17089946
23. Mosnier A, Caini S, Daviaud I, Bensoussan JL, Stoll-Keller F, Bui TT, et al. GROG network. Ten influenza seasons in France: distribution and timing of influenza A and B circulation, 2003-2013. *BMC Infect Dis*. 2015;15(1):357. <https://doi.org/10.1186/s12879-015-1056-z> PMID: 26289794
24. Vergu E, Grais RF, Sarter H, Fagot JP, Lambert B, Valleron AJ, et al. Medication sales and syndromic surveillance, France. *Emerg Infect Dis*. 2006;12(3):416-21. <https://doi.org/10.3201/eid1203.050573> PMID: 16704778
25. European Centre for Disease Prevention and Control (ECDC). Influenza case definitions. 2016. [Accessed 6 Sep 2016]. Stockholm: ECDC. Available from: http://ecdc.europa.eu/en/healthtopics/influenza/surveillance/Pages/influenza_case_definitions.asp
26. Bayer C, Remschmidt C, an der Heiden M, Tolksdorf K, Herzhoff M, Kaersten S, et al. Internet-based syndromic monitoring of acute respiratory illness in the general population of Germany, weeks 35/2011 to 34/2012. *Euro Surveill*. 2014;19(4):20684. <https://doi.org/10.2807/1560-7917.ES2014.19.4.20684> PMID: 24507468
27. Core R. Teal and contributors worldwide. The R Stats Package. Available from: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/filter.html>.
28. Vega T, Lozano JE, Meerhoff T, Snacken R, Beauté J, Jorgensen P, et al. Influenza surveillance in Europe: comparing intensity levels calculated using the moving epidemic method. *Influenza Other Respi Viruses*. 2015;9(5):234-46. <https://doi.org/10.1111/irv.12330> PMID: 26031655
29. Wenham C, Gray ER, Keane CE, Donati M, Paolotti D, Pebody R, et al. Self-Swabbing for Virological Confirmation of Influenza-Like Illness Among an Internet-Based Cohort in the UK During the 2014-2015 Flu Season: Pilot Study. *J Med Internet Res*. 2018;20(3):e71. <https://doi.org/10.2196/jmir.9084> PMID: 29496658
30. Jackson ML, Nguyen M, Kirlin B, Madziwa L. Self-Collected Nasal Swabs for Respiratory Virus Surveillance. *Open Forum Infect Dis*. 2015;2(4):ofv152. <https://doi.org/10.1093/ofid/ofv152> PMID: 26613095
31. Goyal S, Prasert K, Praphasiri P, Chittaganpitch M, Waicharoen S, Ditsungnoen D, et al. The acceptability and validity of self-collected nasal swabs for detection of influenza virus infection among older adults in Thailand. *Influenza Other Respi Viruses*. 2017;11(5):412-7. <https://doi.org/10.1111/irv.12471> PMID: 28741903
32. World Health Organization (WHO). WHO Regional Office for Europe guidance for sentinel influenza surveillance in humans. 2011. Geneva: WHO. [Accessed 15 Apr 2015]. Available from: http://www.euro.who.int/__data/assets/pdf_file/0020/90443/E92738.pdf?ua=1
33. Rehn M, Carnahan A, Merk H, Kühlmann-Berenzon S, Galanis I, Linde A, et al. Evaluation of an Internet-based monitoring system for influenza-like illness in Sweden. *PLoS One*. 2014;9(5):e96740. <https://doi.org/10.1371/journal.pone.0096740> PMID: 24824806

License and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence, and indicate if changes were made.

This article is copyright of the authors, 2018.

4.4. Perspectives

Ce travail a permis la validation d'une méthode de calcul de signaux observés par le biais d'une cohorte participative. La méthode est désormais appliquée en routine dans le calcul des incidences hebdomadaires des participants atteints d'une maladie grippale. Son extension à d'autres pathologies infectieuses dont l'estimation peut être faite à partir des symptômes recueillis (comme les gastroentérites) nécessite encore d'être testée. Si elle est validée, la cohorte participative GrippeNet.fr permettra le suivi en population de plusieurs pathologies de façon standardisée. Cette extension pourrait également être testée et appliquée au niveau européen pour standardiser les analyses des signaux recueillis dans la plateforme InfluenzaNet. Une application à partir des incidences recueillies en temps réel est prévue pour comprendre plus en détails l'évolution du signal depuis son estimation en temps réel jusqu'à la consolidation des données (i.e. les données sur lesquelles nous avons basé notre présente étude).

Puisque la définition de l'ECDC suit plus difficilement le syndrome grippal que l'IRA dans la population, et grâce à la flexibilité du système, la cohorte française pourrait tout à fait utiliser deux définitions pour suivre à la fois les syndromes grippaux et les IRA. Cela permettrait d'un côté d'utiliser une définition plus spécifique pour comparer l'incidence des syndromes grippaux détectés à celle du réseau Sentinelles. De l'autre, cela permettrait de continuer à avoir des résultats et un suivi commun avec les autres pays partenaires d'InfluenzaNet.

Si la tendance et l'évolution temporelle des signaux de GrippeNet.fr, en comparaison à ceux des systèmes traditionnels, nous conforte dans sa capacité à suivre les maladies grippales avec une grande précision, l'estimation de l'intensité propre au système est toujours à étudier. Pour comprendre cette intensité, il est nécessaire d'avoir des données historiques sur lesquelles baser les comparaisons. Cela est désormais possible grâce aux huit saisons de données disponibles, ce qui n'était pas le cas au moment de notre étude. Nous pouvons alors mettre en place une méthode standardisée d'estimation d'un seuil épidémique en population générale, qui permettrait de définir différents paliers d'intensité. Pour ce faire, différentes méthodes statistiques existent comme la régression périodique, dites « Serfling », utilisée par le réseau Sentinelles pour détecter les épidémies de grippe et de gastroentérites [41]. En Europe, la méthode MEM (Moving Epidemic Method) a été adoptée comme méthode standard de classification de l'activité grippale à partir des données de soins primaires [187]. La comparaison de l'intensité entre les saisons ou les pays pour comprendre les tendances des épidémies saisonnières est essentielle [188]. La MEM a aussi été appliquée pour l'analyse de sources de données issues de la surveillance syndromique [189]. Ainsi, l'analyse standardisée des données de GrippeNet.fr et des autres pays d'InfluenzaNet à partir de cette méthode paraît appropriée. A partir d'un jeu de données de cinq années de surveillance minimum, la MEM calcule le moment et la durée d'une épidémie de grippe en déterminant le nombre minimal de semaines et le taux cumulatif maximal [187, 188]. Les taux d'activité grippale qui surviennent de part et d'autre de ces périodes épidémiques sont définis comme activité de base (i.e. activité pré- et post-

épidémique) et servent à calculer le seuil épidémique. Lorsque les incidences des maladies grippales franchissent le seuil épidémique, c'est le début d'une activité grippale importante dans la population suivie. Cinq niveaux d'activité sont alors recensés : de base, faible, moyen, élevé, et très élevé.

Enfin, la capacité de GrippeNet.fr à suivre plus finement les IRA est un atout qu'il serait néanmoins important de valoriser. Le fardeau que représente le VRS dans la population est considérable, avec plus de 3 millions d'hospitalisations estimées dans le monde, impactant surtout les nourrissons et les personnes très âgées [190]. Cependant, le rôle du VRS dans l'apparition de maladies chez les adultes souffrant de problèmes médicaux chroniques et chez les personnes très âgées est mal défini [190]. En 2015, le comité consultatif pour la production et le développement des vaccins de l'OMS (PDVAC) a mis en évidence l'importance du développement d'un vaccin contre le VRS compte tenu de son impact sur la morbi-mortalité [190]. Pour soutenir ces développements, il est nécessaire de rassembler des données épidémiologiques et virologiques sur le VRS, afin de mieux connaître le fardeau qu'il représente, sa saisonnalité et ses caractéristiques virologiques. En France, depuis la saison 2016/2017, le réseau Sentinelles a intégré la surveillance des IRA chez les personnes de 65 ans et plus, en plus de la surveillance traditionnelle des syndromes grippaux. D'autres dispositifs se développent, notamment dans les EHPAD, pour mieux caractériser l'épidémiologie liée à ce virus chez les personnes âgées voire très âgées. De par la finesse des données collectées, GrippeNet.fr pourrait être un bon candidat à la mise en place d'une surveillance des IRA en population générale, pour compléter les récentes surveillances mises en place. Cela permettrait de répondre à certaines problématiques inhérentes en population générale, comme le fardeau et la saisonnalité des virus occasionnant les IRA. Pour ce faire, une surveillance complémentaire virologique à partir de la cohorte pourrait tout à fait s'envisager. En effet, nos partenaires britanniques ont mis en place une surveillance virologique lors de la saison grippale 2014/2015 pour caractériser les virus respiratoires hivernaux présents chez les participants avec un syndrome grippal [182]. Un suivi virologique a également été mis en place dans la cohorte participative allemande [183].

Les messages clés du Chapitre 4

- La surveillance participative engendre de nouveaux enjeux à estimer, comme la capacité de la population à évaluer son état de santé.
- Les systèmes traditionnels ont permis de valider notre système sur le plan de la tendance et de l'évolution du signal détecté. La surveillance participative semble plus encline à suivre l'infection respiratoire aiguë.
- L'intensité du signal suivi est propre au système participatif. Un historique de minimum cinq saisons est nécessaire pour estimer un seuil épidémique en population générale.
- Ce travail a permis la définition d'un cadre standardisé pour l'estimation des incidences des maladies grippales à partir d'une cohorte participative.

Chapitre 5.

Facteurs associés aux syndromes grippaux à partir d'une cohorte participative en population générale de 2012/2013 à 2017/2018

L'analyse des facteurs associés à la grippe repose essentiellement sur la collecte de données recueillies auprès de professionnels de santé, et plus particulièrement à partir de services hospitaliers [191]. Les facteurs de risque ainsi observés sont donc souvent signe de grippe sévère voire grave. Leur étude permet de mieux connaître les caractéristiques des personnes malades ayant recours à des soins de santé. Les facteurs fréquemment retrouvés comme étant responsables de grippe sévère sont un âge élevé et la présence de comorbidités [192]. A partir de ces informations, les autorités sanitaires peuvent alors établir des recommandations ciblées, comme les recommandations vaccinales antigrippales, pour réduire l'impact de la maladie chez ces sujets.

Cependant, les infections dues à la grippe sont souvent bénignes et ne nécessitent pas le recours à des professionnels de santé. Des études en population sont donc menées en parallèle, et notamment au sein de foyers familiaux, afin de mieux comprendre les mécanismes de transmission de la grippe. Elles ont permis d'identifier des facteurs de transmission liés à l'âge, à la composition de la famille, aux souches virales circulantes, à l'exposition ou encore à l'immunité antérieure [193]. Ces études sont d'autant plus utiles car environ 30% des infections grippales sont transmises à domicile [194]. Malgré tout, les facteurs de risque identifiés ne peuvent pas être extrapolables à la population générale car ces études sont souvent basées à partir d'un cas index engendrant des cas d'infection secondaire [195]. Il a en effet été montré qu'être membre d'un ménage contenant un cas de grippe avéré représente le plus grand facteur de risque d'infection dans le foyer [194, 196]. D'autre part, le nombre de déterminants étudiés reste limité et repose principalement sur les caractéristiques sociodémographiques, les comportements de contact ou l'efficacité des mesures de prévention [197, 198]. Les autres études menées en population générale sont des études de cohortes mises en place lors de la pandémie grippale de 2009, avec l'émergence de la nouvelle souche virale A(H1N1)pdm09. Leurs objectifs étaient de mieux caractériser le virus émergent et son impact sur la population [199, 200]. Même si ces analyses permettent de mieux connaître les facteurs de risque d'infection, ils restent spécifiques au virus pandémique et peuvent potentiellement varier pour les virus saisonniers, en impactant différemment la population [201]. En effet, lors de la pandémie grippale, les personnes les plus à

risque d'infection et de sévérité étaient plus jeunes que les personnes habituellement infectées lors des épidémies saisonnières [202, 203].

Ainsi, l'étude des facteurs de risque d'infection grippale directement à partir de la population générale et lors d'épidémies saisonnières permettrait de mieux cibler les mesures de prévention et de contrôle existantes, et limiter la transmission dans la population et donc a fortiori chez les personnes à risque de complications.

5.1. Résumé de l'article

Chaque épidémie de grippe a un coût sociétal important, tant sur le plan sanitaire (morbidité, mortalité) que sur le plan économique (coûts médicaux directs, perte de productivité). Les manifestations cliniques de la grippe se présentent sous la forme de symptômes généraux et respiratoires, communs à différents virus respiratoires. Sa transmission s'effectue essentiellement par voie respiratoire, de personne à personne. Grâce aux données épidémiologiques fines collectées par la cohorte participative GrippeNet.fr, nous pouvons mettre en place des analyses plus spécifiques à un niveau individuel, en plus du suivi des incidences des maladies grippales effectué à un niveau populationnel. La connaissance des facteurs associés aux maladies grippales au cours des épidémies saisonnières en population générale est alors possible.

Pour mettre en évidence ces déterminants, nous avons analysé les données épidémiologiques collectées via GrippeNet.fr entre 2012/2013 et 2017/2018, et liées aux caractéristiques sociodémographiques et sanitaires des participants, mais aussi à leur mode de vie. La connaissance hebdomadaire de l'état de santé des participants nous a permis de déterminer les facteurs associés à au moins un épisode de maladie grippal par saison de surveillance. Des études similaires ont été conduites via les autres plateformes européennes [139, 144, 166] et en France pour la saison 2015/2016 [139]. Nous avons décidé de reconduire ce travail en France et sur plusieurs saisons car plusieurs limites apparaissaient sur les travaux précédemment menés. Le seul travail européen qui ait analysé les facteurs associés aux syndromes grippaux sur plusieurs saisons n'a pas inclus la France à l'analyse. Les déterminants pourraient être spécifiques d'un pays à l'autre, car des différences culturelles peuvent exister. Par ailleurs, ce travail avait fait l'hypothèse de l'indépendance des participants d'une saison à l'autre [166]. Or, dans la cohorte française, près de 80% des participants d'une saison donnée sont des anciens participants, cette hypothèse ne peut donc se vérifier. L'étude menée sur une seule saison en France comportait un nombre plus faible de participants pouvant rendre les estimations moins robustes qu'une estimation faite sur plusieurs saisons et en prenant en compte des caractéristiques propres à notre cohorte comme la non-indépendance des participants.

Selon les saisons, de 11% (2013/2014 ; n= 469/4140) à 29% (2012/2013 ; n= 866/2943) des participants ont eu au moins un épisode grippal durant une épidémie saisonnière. Les facteurs associés à un risque plus élevé d'épisode grippal étaient : le sexe féminin, un jeune âge (< 5 ans,

et entre 5-14 ans), les allergies respiratoires, recevoir un traitement pour une maladie chronique, le surpoids et l'obésité, utiliser les transports publics et avoir un contact avec un animal domestique. Un âge avancé et le fait d'être vacciné contre la grippe étaient des facteurs protecteurs d'un épisode grippal. Grâce à la flexibilité du système, nous avons pu estimer les facteurs associés aux syndromes grippaux en utilisant différentes définitions plus ou moins spécifiques. Les facteurs retrouvés dans notre étude correspondent à ceux présentés dans d'autres études s'appuyant essentiellement sur les professionnels de santé. Le niveau de recueil des données nous a permis aussi de retrouver des facteurs moins établis dans la littérature (comme le fait d'être une femme ou l'utilisation des transports en commun), ou de compléter les facteurs retrouvés (i.e. vivre avec un animal domestique).

Cette étude valide encore un peu plus l'utilisation de systèmes participatifs directement à partir de la population générale, cette fois à une échelle individuelle. Elle confirme ici que les cohortes participatives sont efficaces pour étudier les déterminants des maladies grippales en population.

5.2. Article : Factors associated with influenza-like-illness : a crowdsourced cohort study from 2012/13 to 2017/18

RESEARCH ARTICLE

Open Access



Factors associated with influenza-like-illness: a crowdsourced cohort study from 2012/13 to 2017/18

Caroline Guerrisi^{1*} , Marie Ecollan^{1,2}, Cécile Souty¹, Louise Rossignol¹, Clément Turbelin¹, Marion Debin¹, Thomas Goronflot¹, Pierre-Yves Boëlle¹, Thomas Hanslik^{1,3,4}, Vittoria Colizza¹ and Thierry Blanchon¹

Abstract

Background: Influenza generates a significant societal impact on morbidity, mortality, and associated costs. The study objective was to identify factors associated with influenza-like-illness (ILI) episodes during seasonal influenza epidemics among the general population.

Methods: A prospective study was conducted with the GrippeNet.fr crowdsourced cohort between 2012/13 and 2017/18. After having completed a yearly profile survey detailing socio-demographic, lifestyle and health characteristics, participants reported weekly data on symptoms. Factors associated with at least one ILI episode per influenza epidemic, using the European Centre for Disease Prevention and Control case definition, were analyzed through a conditional logistic regression model.

Results: From 2012/13 to 2017/18, 6992 individuals participated at least once, and 61% of them were women ($n = 4258$). From 11% ($n = 469/4140$ in 2013/14) to 29% ($n = 866/2943$ in 2012/13) of individuals experienced at least one ILI during an influenza epidemic. Factors associated with higher risk for ILI were: gender female (OR = 1.29, 95%CI [1.20; 1.40]), young age (< 5 years old: 3.12 [2.05; 4.68]); from 5 to 14 years old: 1.53 [1.17; 2.00]), respiratory allergies (1.27 [1.18; 1.37]), receiving a treatment for chronic disease (1.20 [1.09; 1.32]), being overweight (1.18 [1.08; 1.29]) or obese (1.28 [1.14; 1.44]), using public transport (1.17 [1.07; 1.29]) and having contact with pets (1.18 [1.09; 1.27]). Older age (≥ 75 years old: 0.70 [0.56; 0.87]) and being vaccinated against influenza (0.91 [0.84; 0.99]) were found to be protective factors for ILI.

Conclusions: This ILI risk factors analysis confirms and further completes the list of factors observed through traditional surveillance systems. It indicates that crowdsourced cohorts are effective to study ILI determinants at the population level. These findings could be used to adapt influenza prevention messages at the population level to reduce the spread of the disease.

Keywords: Human influenza, Epidemiology, Crowdsourcing, Cohort studies

Background

Seasonal influenza represents a major cause of morbidity and mortality worldwide, responsible for 3 to 5 million of serious illnesses [1], and for 290,000 to 650,000 deaths annually, according to recent updates from the World Health Organization and the United States Centers for Disease Control and Prevention [2]. Clinical manifestations occur through influenza-like-illness (ILI)

with sudden onset of fever, myalgia and respiratory signs [3].

Documented influenza risk factors are related to (i) individual characteristics, such as age (higher risk of infection for young age, higher risk of complication and mortality for older age) [4, 5], immunodeficiency [1], pregnancy [6], chronic underlying medical conditions and respiratory diseases [7]; (ii) individual's household features, such as living with children [8]; or (iii) individual's profession like having contacts with children [9] or infected individuals [10]. However, the number of factors analyzed per study is often limited and mainly identified through

* Correspondence: caroline.guerrisi@upmc.fr

¹Sorbonne Université, INSERM, Institut Pierre Louis d'Épidémiologie et de Santé Publique (IPLESP), F-75012 Paris, France

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



traditional surveillance systems based on healthcare professionals [6, 7] or households follow-up [5, 8, 10].

Risk factors identified through healthcare systems or household studies may not be generalized to the general population, as they pertain more severe cases, or individuals selected based on their influenza susceptibility. Exploring risk factors for influenza directly from the general population can help to have a better knowledge of a larger spectrum of infections, including milder infections. Targeting message to individuals at risk of influenza and not only to those at risk of severe influenza may contribute to help limiting the spread of the disease in the population and potentially reducing the associated costs. An influenza epidemic in France is estimated to cost around \$2.6 billion, with \$0.3 billion of direct costs of medical care and \$2.3 billion of indirect costs due to loss of productivity, leading to 2.9 (± 2.5) days of work lost per influenza episode and person [11, 12].

Risk factors analyses based on the general population have been already addressed, but most were implemented during the 2009 A/H1N1pdm09 pandemic [13]. Since 2009, a participatory syndromic surveillance system for influenza, called Influenzanet, is operational in Europe [14, 15]. The system allows fine scale data collection among the general population enabling detailed risk factors analyses [16]. Here we focus on six consecutive influenza seasons, from 2012/13 to 2017/18, to estimate ILI frequency among the GrippeNet.fr (GN) cohort in France and identify the factors associated with ILI infections.

Methods

GrippeNet.fr data collection

GN is a crowdsourced surveillance system operating each year from November to April in mainland France since 2012. It is part of a broader European platform Influenzanet, where ten countries are involved [15]. Individuals from the GN cohort report their influenza-related symptoms through a dedicated website (<https://www.grippenet.fr>). Consent is informed and implied through registration and voluntarily completing a profile survey. This profile survey can be updated throughout the season, and covers socio-demographic (gender, age, household composition, occupation, place of residency); lifestyle (having pets, daily contacts, daily transportation means, smoking habit); and health-related characteristics (height and weight to estimate the body mass index (BMI), chronic treatments for at least one comorbidity including asthma, diabetes, immunosuppression, heart, kidney or pulmonary diseases, influenza vaccination status for the current and past seasons, respiratory allergies) [17, 18]. After the profile completion, symptoms data are collected on a weekly basis [14–16]. If symptoms are reported (from a list of 19 symptoms), further

questions are asked to detail them and the participant behavior [19]. The profile and weekly symptoms surveys were published elsewhere (profile survey [17] and symptoms survey [19]). Participants can also answer to profile surveys and weekly symptoms surveys for other household members through multi-user accounts to facilitate for example participation and report of children and elderly. Each participant added in this way has all her/his individual information filled in the platform and was considered as an individual participant in our analyses.

Although GN participants were not representative of the French general population in terms of age and gender (overrepresentation of middle-aged individuals and women), all age classes were represented (data from 2011/12 season [17]). The GN population was also found to be more frequently employed, with a higher education level. No significant difference was found regarding chronic conditions, such as asthma and diabetes.

Study period

The analysis was conducted on six GN seasons from 2012/13 to 2017/18.

Study participants

Each season, GN participants who provided regular information were included in the current study. Regular information was defined as having filled at least one profile survey and three weekly symptoms surveys for a given season. At least one symptoms survey should have been completed before, another during and one after the influenza epidemic period as stated by the French national surveillance network in primary care, called *Sentinelles* network [20]. A similar inclusion criterion was used in a previous work conducted at the European level [16]. Individuals included in the study may have participated in one to six GN seasons.

ILI case definitions

Different case definitions have been built from symptoms declared on GN [19]. For the risk factors analyses, we considered the ILI definition of the European Centre for Disease Control and Prevention (ILI_{ECD}), which is often used by other Influenzanet studies [16, 21, 22]. ILI is defined as (i) the sudden onset of symptoms, AND (ii) at least one of the following four systemic symptoms (fever or feverishness, malaise, headache or myalgia), AND (iii) at least one of the three respiratory symptoms (cough, sore throat, shortness of breath) [23]. The analyses were also conducted on a much more specific definition, closer to the one used by the French general practitioners (GP) *Sentinelles* network, stated as GN_{ILI} and defined by (i) sudden onset of symptoms, AND (ii) fever $\geq 38^\circ\text{C}$ or fever (when the body temperature level

was not available) AND pain or headache, AND (iii) sore throat or cough or shortness of breath [19]. The latter was computed as a sensitivity analysis, in order to clarify the ILI risk factors found.

As demonstrated in previous research, ILI represents a good proxy for influenza estimates [3, 24]. As no influenza virological confirmation was available, we took only into account ILI occurring during the influenza epidemic period identified by the *Sentinelles* network.

Statistical analyses

Characteristics of the study participants between 2012/13 to 2017/18 were described depending on the last profile survey completed over the study period. As GN is an observational cohort, most of the participants are returning participants season after season, and so one participant can be counted in the study from one to six times. A person-season was defined as an individual having participated in the study during one season. Determinants associated with having at least one ILI episode during a follow-up season were estimated through a conditional logistic regression model, using the generalized estimated equations for longitudinal correlated data, based on the person-seasons. Explanatory variables (season, socio-demographic, geographic, health-related and lifestyle characteristics) were tested in univariate analyses. All covariates with a *p*-value below 0.2 were tested in multivariate analyses. Covariates were selected through a backward stepwise selection. The final model included all covariates associated with having at least one ILI episode with a *p*-value below or equal to 0.05. Same analyses were conducted using the more specific

definition (GN_{ILI}). All statistical analyses were performed using the R software (3.2.5 version) and the *gee-pack* package [25, 26].

Results

Participation

Participation description is available in Table 1. Overall, there were 6992 individuals who participated at least once over the study period from 2012/13 to 2017/18. The 6992 individuals represented 24,653 person-seasons over the 6 year follow-up. An individual participated on average three seasons from 2012/13 to 2017/18. The mean number of participants per season was 4109 (min = 2943 in 2012/13 and max = 4780 in 2015/16). Individuals filled a mean number of 19 weekly symptoms surveys per season over the 6 years follow-up, with seasons lasting from 19 to 23 weeks.

Description of the participants

Socio-demographic, lifestyle and health-related characteristics of the 6992 participants included in the study were described in Table 2. Overall, participants were mostly women (61%, *n* = 4258), living in urban area (81%, *n* = 5637), and having a mean age of 51 years old. Regarding daily contacts, 32% (*n* = 2268) were in contact with groups of ≥10 individuals, 26% (*n* = 1788) with children (beyond their own ones), 10% (*n* = 728) with patients and 10% (*n* = 721) with elderly. Vaccination against influenza for the current season was done in 36% (*n* = 2517) of the individuals followed (min = 29% in 2013/14, and max = 42% in 2017/18). Concerning the underlying health conditions, 22% (*n* = 1504) of the participants were treated for at least one comorbidity

Table 1 Participation and influenza-like-illness (ILI) cases reported per season from 2012/13 to 2017/18, GrippeNet.fr

	2012/13	2013/14	2014/15	2015/16	2016/17	2017/18
Epidemic periods (starting week-ending week; number of epidemic weeks) ^a	2012w51-2013w11 (13 weeks)	2014w05-2014w09 (5 weeks)	2015w03-2015w11 (9 weeks)	2016w04-2016w14 (11 weeks)	2016w50-2017w05 (8 weeks)	2017w50-2018w10 (13 weeks)
GN registered individuals ^b , N	6046	5907	6632	6515	6234	5958
GN participants included in the study ^c , n (%)	2943 (49%)	4140 (70%)	4428 (67%)	4780 (73%)	4204 (67%)	4158 (70%)
Symptoms surveys filled in by the included participants, N	58,235	76,578	80,106	99,333	73,711	82,252
Included participants having at least one ILI episode during the epidemic period (ILLeCDC definition), n (%)	866 (29%)	469 (11%)	886 (20%)	829 (17%)	787 (19%)	855 (21%)
Included participants having at least one ILI episode during the epidemic period (GN _{ILI} definition), n (%)	231 (8%)	111 (3%)	251 (6%)	203 (4%)	174 (4%)	198 (5%)

GN GrippeNet.fr

^aEpidemic periods defined by the French *Sentinelles* network (available on <http://www.sentiweb.fr/?page=epidemies>)

^bAny GN participants having filled in at least one profile survey

^cAny GN participants having filled in at least three symptoms surveys (one before, one during and one after the influenza epidemic period)

Table 2 Socio-demographic, exposure and health characteristics of GrippeNet.fr participants over the 2012/13 to 2017/18 study period and demographic characteristics of the French metropolitan population (Insee^a, 2015 data). (For comprehensive results of the representativeness of GrippeNet.fr participants, see ref. [17])

	2012/13–2017/18 study period ^b N = 6992	Metropolitan France general population (Insee ^a , 2015 data) N = 64,604,773
Sociodemographic characteristics		
Gender (m.d. = 0)		
Female	4258 (61%)	52%
Male	2734 (39%)	48%
Age (m.d. = 10)		
[0–5]	109 (2%)	6%
[5–15]	384 (5%)	12%
[15–45]	1900 (27%)	37%
[45–65]	2577 (37%)	26%
[65–75]	1567 (22%)	10%
75	445 (6%)	9%
Household composition (m.d. = 16)		
Living alone	1063 (15%)	36%
Living with ≥1 child	2435 (35%)	32%
Living with adults only	3478 (50%)	32%
Occupation (m.d. = 123)		
Working	3345 (49%)	40%
Student	774 (11%)	24%
Unemployed	160 (2%)	10%
Retired	2321 (34%)	25%
Stay at home/Sick leave	266 (4%)	–
Place of residency (m.d. = 0)		
Urban	5637 (81%)	77%
Rural	1355 (19%)	23%
Lifestyle		
Use of public transport (m.d. = 0)	1169 (17%)	–
Individuals' contacts (m.d. = 0)		
Contact with patients	728 (10%)	–
Contact with elderly	721 (10%)	–
Contact with a group of people (≥10)	2268 (32%)	–
Contact with children	1788 (26%)	–
At least one pet at home (m.d. = 14)	3199 (46%)	–
Health characteristics		
Influenza vaccination (current season) (m.d. = 6)	2517 (36%)	–
Smokers (m.d. = 7)	766 (11%)	–
At least one health comorbidity ^c (m.d. = 0)	1504 (22%)	–

Table 2 Socio-demographic, exposure and health characteristics of GrippeNet.fr participants over the 2012/13 to 2017/18 study period and demographic characteristics of the French metropolitan population (Insee^a, 2015 data). (For comprehensive results of the representativeness of GrippeNet.fr participants, see ref. [17]) (Continued)

	2012/13–2017/18 study period ^b N = 6992	Metropolitan France general population (Insee ^a , 2015 data) N = 64,604,773
Chronic treatment^c for: (m.d. = 0)		
Asthma	441 (6%)	–
Diabetes	263 (4%)	–
Heart diseases	686 (10%)	–
Kidney diseases	50 (1%)	–
Immunosuppression	210 (3%)	–
Pulmonary diseases	195 (3%)	–
At least one respiratory allergy (m.d. = 0)	2413 (35%)	–
BMI (m.d. = 139)		
Underweight	348 (5%)	–
Normal weight	3946 (58%)	–
Overweight	1799 (26%)	–
Obese	760 (11%)	–

^aThe National Institute of Statistics and Economic Studies (<https://www.insee.fr/fr/accueil>) collects, analyses and disseminates information on the French economy and society

^bParticipants included in the analysis may have participated for the six GrippeNet.fr seasons. An individual may have participated up to six times (one time for each followed year). Here we analyze the participants' characteristics depending on the last survey completed over the study period

^cParticipants receiving a chronic treatment for at least one of the following diseases: asthma, diabetes, immunosuppression, heart, kidney, and pulmonary diseases
m.d. missing data. The number corresponds to the total of missing data for a given variable among the 6992 participants

(as known as asthma, diabetes, immunosuppression, heart, kidney and pulmonary diseases). One third ($n = 2413$) had respiratory allergies and 11% ($n = 766$) were smokers. Further details are available in Table 2. The description of the participants for the six seasons from 2012/13 to 2017/18 did not show any difference between seasons. Full results are detailed in Additional file 1.

Estimations of ILI episodes

Depending on GN season, from 11% (469/4140 in 2013/14) to 29% (866/2943 in 2012/13) of individuals experienced at least one ILI_{ECDC} episode during the influenza epidemic. Using the more specific GN_{ILI} definition, this range reduced from 3% (111/4140 in 2013/14) to 8% (231/2943 in 2012/13) (Table 1).

Risk factor analyses: univariate and multivariate analyses

In the univariate analysis, factors associated with having at least one ILI_{ECDC} episode during the influenza

epidemic period with a p -value below 0.2 were: gender, age, household composition, occupation, use of public transport, pets at home, contacts with patients, contact with a group of individuals, contact with children, influenza vaccination status regarding the current and past seasons, being treated for health comorbidities, respiratory allergy, and body mass index (Table 3).

In the final multivariate model (Table 3), women tended to have more ILI_{ECDC} episodes with respect to men (OR = 1.29, [1.20; 1.40]). Children experienced more ILI_{ECDC} episodes compared to adults from 45 to 64 (OR = 3.12 [2.05; 4.68] for children from 0 to 4, and OR = 1.53 [1.17; 2.00] for children from 5 to 14). Individuals using public transport (OR = 1.17 [1.07; 1.29]), having pets at home (OR = 1.18 [1.09; 1.27]), treated for health comorbidities (OR = 1.20 [1.09; 1.32]), having respiratory allergy (OR = 1.27 [1.18; 1.37]), or being overweight or obese (OR = 1.18 [1.08; 1.29] and OR = 1.28 [1.14; 1.44], respectively) were also more prone to experience an ILI_{ECDC} episode. On the contrary, individuals aged 75 years or older tended to have fewer ILI_{ECDC} episodes compared to those in the range of 45 to 64 (OR = 0.70 [0.56; 0.87]). Individuals vaccinated for the current season were also found to be at lower risk (OR = 0.91 [0.84; 0.99]).

The sensitivity analysis conducted with the more specific GN_{ILL} definition showed similar results (see Additional file 2).

Discussion

This study allowed us to identify factors associated with ILI directly from the general population. Some health determinants already described in the literature were found, such as having a young age or health comorbidities. We also found more debated risk factors, such as sex and the use of public transport, or rather unexpected ones, such as living with pets.

In our analyses, women were found to be at higher risk for ILI, independently of the case definition used (either ILI_{ECDC} or GN_{ILL} definitions). Female vs. male differences regarding influenza have been previously evaluated: at younger (< 20 years) and older (> 80) ages, morbidity rates seem to be higher for males than females, however during the reproductive age (from 20 to 49 age group) women were found to have higher morbidity rates [27–29]. The observed differences are thought to be based on various factors that can affect both sex (e.g. genetic, immunological, and hormonal differences) and gender (i.e. behavioural) characteristics [27, 29, 30]. Here, the increased risk for women is found even adjusting for behavioural determinants such as living or having contacts with children. Further investigations should be implemented to understand better the biological and behavioural impacts.

Taking public transport is associated in our study with an increased risk for ILI. Only few studies addressed this association. One found that public transport used within 5 days of symptoms onset was associated with an increased risk of consulting for acute respiratory infection [31]. Previous Influenzanet works did not find an association between public transport and influenza [16, 21, 32], likely because of lack of statistical power due to the consideration of one season only [16, 21] or due to methodological differences based on the public transport covariate definition [32]. In this last article, the public transport covariate was defined using three categories (bicycle/foot, car and public transport), whereas here we opted for two categories (private vs. public transports) in order to better observe any impact of public transports with respect to other modes of daily locomotion, where individuals do not have close contacts.

Individuals who had pets at home had a higher chance of experiencing ILI as well. This result confirms previous findings obtained in the Influenzanet platform [16, 32], though it still remains unexplained. It would be interesting to further explore additional factors that can impact this small increased risk observed, such as the lifestyle of individuals living with pets and the contacts they establish.

In addition to the risk factor analysis, we estimated here the average fraction of individuals presenting at least one ILI episode during a season, ranging from 5% (GN_{ILL}) to 19% (ILI_{ECDC}). International studies have provided estimates in the same ballpark (2.8 to 10.9% in the US [33], 10 to 25% in Canada [34]), however the comparison can be difficult as influenza impact depends on the specific season but also on the surveillance system and the case definition. In France, an average 3.4% of the population was estimated to have an ILI requiring a medical consultation from general practitioners surveillance data [20], a lower value compared to our GN estimate that can be explained by the limited fraction of illnesses leading to health-seeking behavior (56.7% of GN_{ILL} episodes, 32.6% with the ILI_{ECDC} definition) [35].

The strength of this study is the identification of ILI determinants in the general population. Fine scale data have been collected through a large panel of individuals during several seasons allowing the evaluation of a wide range of ILI determinants. Little-known risk factors for ILI were identified in addition to well-known factors indicating that GrippeNet.fr is effective to study ILI determinants at the population level. However, a few limitations can be highlighted. First, the cohort was not representative of the French population as it is a crowd-sourced system. Nevertheless, all ages and gender were represented. Second, no virological confirmation was

Table 3 Factors associated with at least one ILI episode (ILI_{ECCDC}) during six influenza epidemic periods, GrippeNet.fr

		N ^a	Cases ^b n (%)	OR [CI 95%] Univariate analyses	p-value	OR [CI 95%] Multivariate analysis	p-value
Season	2012/13	2943	866 (29%)	Ref.	<10 ⁻⁴	Ref.	<10 ⁻⁴
	2013/14	4140	469 (11%)	0.31 [0.27;0.34]		0.30 [0.26;0.33]	
	2014/15	4428	886 (20%)	0.59 [0.54;0.66]		0.60 [0.54;0.67]	
	2015/16	4780	829 (17%)	0.50 [0.45;0.55]		0.50 [0.45;0.56]	
	2016/17	4204	787 (19%)	0.54 [0.49;0.60]		0.57 [0.51;0.63]	
	2017/18	4158	855 (21%)	0.61 [0.55;0.68]		0.64 [0.57;0.71]	
Sociodemographic characteristics							
Gender	Male	9742	1623 (17%)	Ref.	<10 ⁻⁴	Ref.	<10 ⁻⁴
	Female	14,911	3069 (21%)	1.30 [1.20;1.40]		1.29 [1.20;1.40]	
Age	[45–65]	9940	1959 (20%)	Ref.	<10 ⁻⁴	Ref.	<10 ⁻⁴
	[0–5]	306	111 (36%)	2.27 [1.78;2.89]		3.12 [2.05;4.68]	
	[5–15]	1133	279 (25%)	1.32 [1.12;1.55]		1.53 [1.17;2.00]	
	[15–45]	5948	1309 (22%)	1.15 [1.06;1.25]		1.13 [1.03;1.25]	
	[65–75]	6003	885 (15%)	0.71 [0.65;0.78]		0.88 [0.77;1.01]	
	≥75	1300	146 (11%)	0.54 [0.45;0.65]		0.70 [0.56;0.87]	
Household composition	Living alone	3917	731 (19%)	Ref.	<10 ⁻⁴	NS	NS
	Living with ≥1 child	7712	1714 (22%)	1.25 [1.12;1.40]			
	Living with adults only	12,953	2234 (17%)	0.92 [0.83;1.02]			
Occupation	Working	11,587	2432 (21%)	Ref.	<10 ⁻⁴	Ref.	<10 ⁻²
	Student	2162	543 (25%)	1.25 [1.11;1.41]		1.02 [0.83;1.26]	
	Unemployed	575	138 (24%)	1.18 [0.96;1.46]		1.18 [0.95;1.46]	
	Retired	8932	1299 (15%)	0.64 [0.59;0.70]		0.80 [0.70;0.91]	
	Stay at home/ Sick leave	993	191 (19%)	0.89 [0.74;1.07]		0.84 [0.70;1.02]	
Place of residency	Rural	4840	906 (19%)	Ref.	0.67	–	–
	Urban	19,813	3786 (19%)	1.02 [0.93;1.12]			
Lifestyle							
Use of public transport	No	20,843	3855 (18%)	Ref.	<10 ⁻⁴	Ref.	<10 ⁻²
	Yes	3810	837 (22%)	1.22 [1.12;1.34]		1.17 [1.07;1.29]	
Pets at home	None	13,492	2377 (17%)	Ref.	<10 ⁻⁴	Ref.	<10 ⁻⁴
	At least one	11,118	2311 (20%)	1.21 [1.13;1.30]		1.18 [1.09;1.27]	
Contacts with patients	No	22,163	4154 (19%)	Ref.	0.01	NS	NS
	Yes	2490	538 (22%)	1.17 [1.04;1.30]			
Contacts with elderly	No	22,153	4234 (19%)	Ref.	0.41	–	–
	Yes	2496	458 (18%)	0.96 [0.86;1.07]			
Contacts with a group of people	No	16,795	3047 (18%)	Ref.	<10 ⁻⁴	NS	NS
	Yes	7858	1645 (21%)	1.17 [1.09;1.25]			
Contacts with children	No	18,722	3342 (18%)	Ref.	<10 ⁻⁴	NS	NS
	Yes	5931	1350 (23%)	1.33 [1.23;1.43]			
Health characteristics							
Influenza vaccination (current season)	No	16,083	3240 (20%)	Ref.	<10 ⁻⁴	Ref.	0.04
	Yes	8554	1450 (17%)	0.82 [0.76;0.88]		0.91 [0.84;0.99]	
Influenza vaccination (last season)	No	15,190	3083 (20%)	Ref.	<10 ⁻⁴	NS	NS
	Yes	9305	1575 (17%)	0.81 [0.75;0.87]			

Table 3 Factors associated with at least one ILI episode (ILI_{ECCDC}) during six influenza epidemic periods, GrippeNet.fr (Continued)

		N ^a	Cases ^b n (%)	OR [CI 95%] Univariate analyses	p-value	OR [CI 95%] Multivariate analysis	p-value
Smoking status	Non smoker	21,995	4176 (19%)	Ref.	0.55	–	–
	Smoker	2630	514 (20%)	1.04 [0.92;1.16]			
Chronic treatment for at least one comorbidity ^c	No comorbidity	19,425	3648 (19%)	Ref.		Ref.	<10 ⁻³
	At least one comorbidity	5228	1044 (20%)	1.08 [0.99;1.17]	0.10	1.20 [1.09;1.32]	
	Asthma	1449	384 (27%)	1.53 [1.33;1.75]	< 10 ⁻⁴		
	Diabetes	941	156 (17%)	0.85 [0.70;1.04]	0.12		
	Heart diseases	2473	420 (17%)	0.88 [0.78;0.99]	0.03		
	Kidney diseases	152	24 (16%)	0.77 [0.48;1.25]	0.29		
	Immunosuppression	655	140 (21%)	1.13 [0.91;1.39]	0.26		
	Pulmonary diseases	635	161 (25%)	1.40 [1.13;1.73]	<10 ⁻²		
Respiratory allergy	None	16,400	2887 (18%)	Ref.	<10 ⁻⁴	Ref.	<10 ⁻⁴
	At least one	8253	1805 (22%)	1.29 [1.20;1.39]		1.27 [1.18;1.37]	
BMI	Normal weight [18.5;25[14,000	2563 (18%)	Ref.	<10 ⁻²	Ref.	<10 ⁻⁴
	Underweight < 18.5	1053	201 (19%)	1.04 [0.88;1.25]		0.92 [0.77;1.11]	
	Overweight [25;30[6571	1241 (19%)	1.04 [0.96;1.13]		1.18 [1.08;1.29]	
	Obese ≥30	2629	586 (22%)	1.26 [1.12;1.40]		1.28 [1.14;1.44]	

OR = Odd Ratio; CI95% = 95% confidence interval; NS = not significant; – = variable not tested in the multivariate analysis as the p-value was higher than 0.2 in univariate analysis

^aN = global number of individuals for the six GrippeNet.fr seasons (from 2012/13 to 2017/18), based on the 24,653 person-seasons. An individual may have participated up to six times (one time for each followed year)

^bThe ILI case definition used here is the ILI_{ECCDC} definition

^cParticipants receiving a chronic treatment for at least one of the following diseases: asthma, diabetes, immunosuppression, heart, kidney, and pulmonary diseases. Inclusion of the gathered variable “At least one comorbidity” only in the analyses.

available to ensure influenza follow-up. Thus, risk factors found here were not specifically associated with influenza infection but with a broader set of respiratory viruses causing similar symptoms. To limit the impact of this aspect, we decided to include only ILI episodes occurring during the influenza epidemic period, and also to explore two ILI case definitions, namely a more sensitive one (ILI_{ECCDC}) and a more specific one (GN_{ILI}).

Conclusion

The identification of risk factors from the general population performed in this study confirms and further completes the list of factors observed through traditional surveillance systems. These findings can help target specific communication and influenza prevention campaigns at the population level aimed at reducing the spread of the disease. Some ILI risk factors, as gender, public transport use and having pets are still debated, they should be further investigated.

Additional files

Additional file 1: Socio-demographic, exposure and health characteristics of GrippeNet.fr participants from seasons 2012/13 to 2017/18. This table represents the description of the participants for each of the six seasons studied. (DOCX 26 kb)

Additional file 2: Factors associated with at least one ILI episode (GN_{ILI}-) during six influenza epidemic periods, GrippeNet.fr. This table highlights the factors associated with an ILI episode using a more specific definition. (DOCX 25 kb)

Abbreviations

BMI: body mass index; ECCDC: European Centre for Disease Prevention and Control; GN: GrippeNet.fr; ILI: influenza-like-illness

Acknowledgements

We thank all current and former participants and members of the GrippeNet.fr system.

Authors' contributions

CG, TB designed and analyzed the results from the prevalence and risk factor analyses. CG, ME, CS, LR, CT, MD, TG, PYB, TH, VC, TB interpreted the results. CG, TB wrote the manuscript. CG, ME, CS, LR, CT, MD, TG, PYB, TH, VC, TB read and approved the final manuscript.

Funding

This work was supported by public funds from Sorbonne Université, Inserm, Santé publique France and the French National Research Agency (ANR). The funding was not specific for the study described in this article. The funder had no role in study design, data collection, data analysis, data interpretation, writing of the report, or in the decision to submit this article for publication. All researchers' decisions have been entirely independent from funders.

Availability of data and materials

GrippeNet.fr databases used in this study are not publicly available, in accordance with the authorization we have from the French National Commission on Informatics and Liberty (CNIL, authorization DR-2012-024).

The datasets used and/or analysed during the current study are available from the authors on reasonable request.

Ethics approval and consent to participate

The consent procedure was informed and implied through the voluntary and anonymous registration of participants on a dedicated website: <https://www.grippenet.fr>, to complete a profile survey. Only an email address is required to participate. Information on confidentiality and data security as well as ethics approval are publicly available on the GrippeNet.fr website (<https://www.grippenet.fr/fr/grippenet/confidentialite-et-securite-des-donnees/>). Even though ethics approval is not required by the French law for epidemiological data collection in the setting of non-interventional biomedical research, GrippeNet.fr was reviewed and approved by the French Advisory Committee for research on information treatment in the field of health (i.e. CCTIRS, authorization 11.565), and by the French National Commission on Informatics and Liberty (i.e. CNIL, authorization DR-2012-024).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Sorbonne Université, INSERM, Institut Pierre Louis d'Épidémiologie et de Santé Publique (IPLESP), F-75012 Paris, France. ²Department of family Medicine, Faculté de Médecine, Université Paris Descartes, Paris, France. ³APHP, Service de Médecine Interne, Hôpital Ambroise-Paré, 92100 Boulogne-Billancourt, France. ⁴UFR des sciences de la santé Simone-Veil, Université de Versailles - Saint-Quentin-en-Yvelines, 78280 Versailles, France.

Received: 16 November 2018 Accepted: 17 June 2019

Published online: 04 July 2019

References

- World health organization: influenza (seasonal) - fact sheet <http://www.who.int/mediacentre/factsheets/fs211/en/> (2018). Accessed 15 Nov 2018.
- Juliano AD, Roguski KM, Chang HH, Muscatello DJ, Palekar R, Tempia S, Cohen C, Gran JM, Schanzer D, Cowling BJ, et al. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet*. 2018;391:1285–300.
- Casalegno JS, Eibach D, Valette M, Enouf V, Daviaud I, Behillil S, Vabret A, Soulayr JC, Benchaib M, Cohen JM, et al. Performance of influenza case definitions for influenza community surveillance: based on the French influenza surveillance network GROG, 2009–2014. *Euro Surveill*. 2017;22:14.
- Fiore AE, Uyeki TM, Broder K, Finelli L, Euler GL, Singleton JA, Iskander JK, Wortley PM, Shay DK, Bresee JS, et al. Prevention and control of influenza with vaccines: recommendations of the advisory committee on immunization practices (ACIP), 2010. *MMWR Recomm Rep*. 2010; 59(RR-8):1–62.
- Riley S, Kwok KO, Wu KM, Ning DY, Cowling BJ, Wu JT, Ho LM, Tsang T, Lo SV, Chu DK, et al. Epidemiological characteristics of 2009 (H1N1) pandemic influenza based on paired sera from a longitudinal community cohort study. *PLoS Med*. 2011;8(6):e1000442.
- Sappenfield E, Jamieson DJ, Kourtis AP. Pregnancy and susceptibility to infectious diseases. *Infect Dis Obstet Gynecol*. 2013;2013:752852.
- Jain S, Kamimoto L, Bramley AM, Schmitz AM, Benoit SR, Louie J, Sugerman DE, Druckenmiller JK, Ritger KA, Chugh R, et al. Hospitalized patients with 2009 H1N1 influenza in the United States, April–June 2009. *N Engl J Med*. 2009;361(20):1935–44.
- Viboud C, Boelle PY, Cauchemez S, Lavenu A, Valleron AJ, Flahault A, Carrat F. Risk factors of influenza transmission in households. *The British journal of general practice: the journal of the Royal College of General Practitioners*. 2004;54(506):684–9.
- Marchbanks TL, Bhattarai A, Fagan RP, Ostroff S, Sodha SV, Moll ME, Lee BY, Chang CC, Ennis B, Britz P, et al. An outbreak of 2009 pandemic influenza A (H1N1) virus infection in an elementary school in Pennsylvania. *Clin Infect Dis*. 2011;52(Suppl 1):S154–60.
- France AM, Jackson M, Schrag S, Lynch M, Zimmerman C, Biggerstaff M, Hadler J. Household transmission of 2009 influenza A (H1N1) virus after a school-based outbreak in new York City, April–May 2009. *J Infect Dis*. 2010; 201(7):984–92.
- Szucs T. The socio-economic burden of influenza. *J Antimicrob Chemother*. 1999;44(Suppl B):11–5.
- Carrat F, Sahler C, Rogez S, Leruez-Ville M, Freymuth F, Le Gales C, Bungener M, Housset B, Nicolas M, Rouzioux C. Influenza burden of illness: estimates from a national prospective survey of household contacts in France. *Arch Intern Med*. 2002;162(16):1842–8.
- Delabre RM, Lapidus N, Salez N, Mansiaux Y, de Lamberterie X, Carrat F. Risk factors of pandemic influenza A/H1N1 in a prospective household cohort in the general population: results from the CoPanFlu-France cohort. *Influenza Other Respir Viruses*. 2015;9(1):43–50.
- Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, Koppeschaar C, Rehn M, Smallegange R, Turbelin C, et al. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. *Clin Microbiol Infect*. 2014;20(1):17–21.
- Influenzanet: <https://www.influenzanet.eu/> (2013). Accessed 25 Apr 2013.
- Guerrisi C, Turbelin C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, Perrotta D, Paolotti D, Smallegange R, Koppeschaar C, et al. Participatory syndromic surveillance of influenza in Europe. *J Infect Dis*. 2016; 214(suppl_4):S386–s392.
- Debin M, Turbelin C, Blanchon T, Bonmarin I, Falchi A, Hanslik T, Levy-Bruhl D, Poletto C, Colizza V. Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France. *PLoS One*. 2013;8(9):e73675.
- Cantarelli P, Debin M, Turbelin C, Poletto C, Blanchon T, Falchi A, Hanslik T, Bonmarin I, Levy-Bruhl D, Micheletti A, et al. The representativeness of a European multi-center network for influenza-like-illness participatory surveillance. *BMC Public Health*. 2014;14:984.
- Guerrisi C, Turbelin C, Souty C, Poletto C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, Colizza V. The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France. *Euro Surveill*. 2018;23:25.
- Souty C, Amoros P, Falchi A, et al. Influenza epidemics observed in primary care from 1984 to 2017 in France: A decrease in epidemic size over time. *Influenza Other Respir Viruses*. 2019;13:148–157. <https://doi.org/10.1111/irv.12620>.
- Adler AJ, Eames KT, Funk S, Edmunds WJ. Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey. *BMC Infect Dis*. 2014;14:232.
- Kjelso CGM, Bang H, Ethelberg S, Grove Krause T. Influmeter - an online tool for self-reporting of influenza-like illness in Denmark. *Infectious Diseases*. 2015;48(4):322–7.
- European Commission. Amending decision 2002/253/EC laying down case definitions for reporting communicable diseases to the community network under decision no 2119/98/EC of the European Parliament and of the council. In: *Official Journal of the European Union*; 2012.
- Aguilera JF, Paget WJ, Mosnier A, Heijnen ML, Uphoff H, van der Velden J, Vega T, Watson JM. Heterogeneous case definitions used for the surveillance of influenza in Europe. *Eur J Epidemiol*. 2003;18(8):751–4.
- Yan J, Fine J. Estimating equations for association structures. *Stat Med*. 2004; 23(6):859–74; discussion 875–857,879–880.
- Halekoh UHS, Yan Y. The R package geeppack for generalized estimating equations. *J Stat Softw*. 2006;15(2).
- World Health Organization: Sex, gender and influenza http://whqlibdoc.who.int/publications/2010/9789241500111_eng.pdf?ua=1 (2010). Accessed 20 Jan 2018.
- Wang XL, Yang L, Chan KH, Chan KP, Cao PH, Lau EH, Peiris JS, Wong CM. Age and sex differences in rates of influenza-associated hospitalizations in Hong Kong. *Am J Epidemiol*. 2015;182(4):335–44.
- Klein SL, Hodgson A, Robinson DP. Mechanisms of sex disparities in influenza pathogenesis. *J Leukoc Biol*. 2012;92(1):67–73.
- vom Steeg LG, Klein SL. Sex matters in infectious disease pathogenesis. *PLoS Pathog*. 2016;12(2):e1005374.
- Troko J, Myles P, Gibson J, Hashim A, Enstone J, Kingdon S, Packham C, Amin S, Hayward A, Nguyen Van-Tam J. Is public transport a risk factor for acute respiratory infection? *BMC Infect Dis*. 2011;11:16.
- van Noort SCC, Koppeschaar C, Van Ranst M, Paolotti D, Gomes G. Ten-year performance of Influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour. *Epidemics*. 2015;13:28–36.
- Rolfes Melissa A, Foppa Ivo M, Shikha G, Brendan F, Lynnette B, Singleton James A, Erin B, Daniel J, Olsen Sonja J, Joseph B, et al. Annual estimates of

the burden of seasonal influenza in the United States: a tool for strengthening influenza surveillance and preparedness. *Influenza Other Respir Viruses*. 2018;12(1):132–7.

34. Santé Canada et l'Agence de santé publique du Canada: La grippe https://www.ctf-fce.ca/Research-Library/Issue2_Article6_FR.pdf Accessed 15 Nov 2018.
35. Ariza M, Guerrisi C, Souty C, Rossignol L, Turbelin C, Hanslik T, Colizza V, Blanchon T. Healthcare-seeking behaviour in case of influenza-like illness in the French general population and factors associated with a GP consultation: an observational prospective study. *BJGP Open*. 2017;1:4.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



5.3. Perspectives

La finesse des données collectées du système GrippeNet.fr a permis la mise en évidence de facteurs encore débattus voire inconnus, ce qui représente un atout supplémentaire des systèmes participatifs dans le cadre de la surveillance globale. Même si les facteurs observés mériteraient d'être précisés par le biais d'analyses spécifiques, ils représentent néanmoins des pistes d'action pour Santé publique France dont les rôles sont entre autres la description et le contrôle des épidémies dans le cadre de la surveillance de la grippe.

Concernant la grippe, des différences entre les hommes et les femmes ont déjà été observées, notamment en ce qui concerne la gravité de la maladie, mais elles restent débattues. Dans notre étude, ce sur-risque ressort pour les femmes même après ajustement sur certains facteurs comportementaux comme le fait de vivre ou d'avoir des contacts avec des enfants. Nous envisageons une nouvelle étude spécifique chez les femmes et par tranches d'âge pour mieux comprendre les facteurs biologiques et comportementaux liés à ce sur-risque. En effet, durant la période de fertilité, les femmes auraient un taux de morbidité supérieur aux hommes [204]. Nous utiliserons différentes sources de données comme les données de soins primaires pour mener ces analyses.

Au-delà de la grippe, des études similaires pourraient être menées à partir d'autres indicateurs de santé comme les gastroentérites ou les allergies respiratoires, afin de préciser les facteurs associés à ces pathologies. Par ailleurs, en cas d'épidémie inhabituelle, la réactivité du système permettrait l'étude en temps réel des facteurs associés à de nouveaux virus émergents, afin de proposer des actions de santé publique adaptées et ciblées. En effet, la gestion de la pandémie grippale de 2009 a montré la nécessité d'avoir des systèmes de détection et d'évaluation précoces et complets de la situation [205]. Un système comme GrippeNet.fr permettrait d'avoir déjà une première vision de ces points en population générale, notamment les groupes d'âge les plus touchés par l'infection, le taux d'attaque clinique ou encore la sévérité de l'infection, qui sont difficiles à estimer en temps réel par les structures de soins [205].

Les messages clés du Chapitre 5

- La surveillance participative via GrippeNet.fr est un système valide et efficace pour l'analyse à l'échelle individuelle de données épidémiologiques fines. En effet, les facteurs de risque des maladies grippales retrouvés ici sont similaires à ceux retrouvés via le système de soins.
- Différents facteurs de risque ont été retrouvés : des facteurs bien connus, débattus ou inconnus jusqu'alors.
- GrippeNet.fr pourrait évaluer les facteurs de risque liés à d'autres pathologies en population, mais aussi en temps réel lors de l'émergence de nouvelles pathologies.

Chapitre 6.

Application de GrippeNet.fr au-delà de la grippe

En parallèle des travaux présentés dans les Chapitres 4 et 5, GrippeNet.fr a permis la mise en place d'études complémentaires dans les champs de la veille sanitaire et de la recherche en santé publique. Des travaux sur la grippe à partir des données collectées en routine ou d'études ancillaires (collecte de données additionnelles) ont pu être menés, comme exposés dans le Chapitre 3. Par ailleurs, la bonne acceptation et l'intérêt de GrippeNet.fr auprès de la population et des acteurs de santé publique ont permis d'utiliser ce système participatif à d'autres fins que le suivi de la grippe en population générale. Des projets de surveillance et de recherche dans le champ des maladies infectieuses plus largement ou de la promotion de la santé ont été et continuent à être réalisés par le biais de la cohorte. Nous allons les détailler dans les sections suivantes.

La Figure 11 expose le nombre des différentes études épidémiologiques menées à partir de GrippeNet.fr et selon le secteur de recherche : analyses réalisées à partir des données collectées en routine sur la grippe (incidences, facteurs de risque, efficacité vaccinale) et sur le système participatif lui-même (participation, satisfaction des participants), ou de données additionnelles dans le champ de la grippe (opinion vaccination antigrippale, recours aux soins, connaissances sur la grippe, suivi de cohortes spécifiques), des maladies infectieuses (gastroentérites) ou de la promotion de la santé (évaluation de mesures de prévention ou de politiques de santé). On dénombre 27 études (publiées, en cours ou en projet) rattachées à la cohorte française, et réalisées soit au niveau national ou européen.

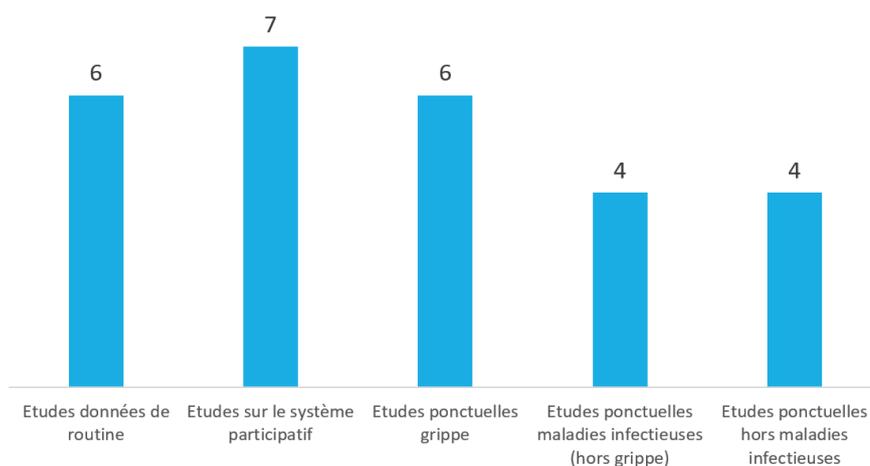


Figure 11. Types et nombre d'études menées avec la cohorte GrippeNet.fr. Parmi les 27 études dénombrées, six sont en cours et deux sont prévues pour la saison 2019/2020.

6.1. GrippeNet.fr pour le suivi d'autres maladies infectieuses

En plus de la grippe, la collecte des symptômes simples chaque semaine permet de suivre d'autres maladies infectieuses directement à partir de la population générale, comme les gastroentérites aiguës ou les allergies respiratoires (Figure 12). Pour les **gastroentérites**, on peut voir une augmentation du signal lors de la dernière semaine de décembre et de la première semaine de janvier (Figure 12a) et pour les **allergies respiratoires**, le signal est plus élevé à la fin de l'hiver, début du printemps que lors des mois de novembre/décembre (Figure 12b). Même si des analyses complémentaires doivent être envisagées, ces signaux sont cohérents par rapport à l'épidémiologie connue des deux pathologies. Les études à mener pourraient être similaires à celles effectuées avec les syndromes grippaux comme la détermination des facteurs associés aux gastroentérites aiguës ou leur suivi temporel durant la période hivernale. Un premier travail en cours de soumission a d'ailleurs étudié les facteurs de risque des gastroentérites aiguës, qui étaient : être en surpoids ou obèse, avoir fréquemment des rhumes ou la grippe, avoir au moins une maladie chronique sous-jacente, vivre seul, et avoir des animaux domestiques. Les facteurs protecteurs retrouvés étaient : être âgé de 10 à 15 ans, ou avoir plus de 60 ans. Un second travail, similaire à celui détaillé au Chapitre 4, est en cours d'étude pour analyser le signal des gastroentérites obtenu via GrippeNet.fr, le comprendre et le comparer aux données issues du système de soins par le biais du réseau Sentinelles.

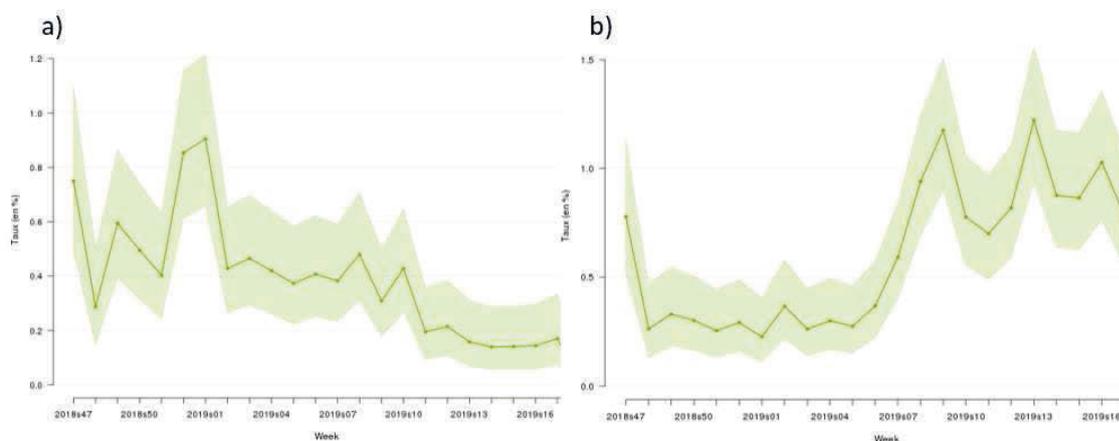


Figure 12. Taux de gastroentérites (a) et d'allergies respiratoires (b) parmi les participants de GrippeNet.fr lors de la saison 2018/2019.

En plus de projets de recherche réalisés à partir des données collectées via GrippeNet.fr, ce sont les caractéristiques propres à la cohorte qui tendent à être exploitées, et notamment sa flexibilité et sa réactivité. En effet, à deux reprises en février 2016 et juin 2018, cette cohorte participative a été utilisée pour **aider à l'investigation d'épidémies de salmonelloses sur le territoire métropolitain**. Santé publique France, en charge de l'investigation de l'épidémie, avec les

responsables de la cohorte GrippeNet.fr, ont mis en place et diffusé auprès des GrippeNauts un questionnaire alimentaire simplifié. L'objectif était d'utiliser les participants de la cohorte GrippeNet.fr comme témoins dans la mise en place d'une étude cas-témoins, pour valider les hypothèses d'investigation établies par SpF. Ces études visaient à identifier les aliments consommés par les cas qui pouvaient différer de ceux consommés par les témoins, dans le but de contenir et de mettre en place les mesures de prévention nécessaires pour éradiquer l'épidémie. Les GrippeNauts ont bien adhéré aux deux enquêtes alimentaires proposées. L'investigation de la première épidémie a fait l'objet d'une publication scientifique [206]. Lors de cette investigation, 2 914 GrippeNauts ont participé, soit 47% des participants de la saison 2015/2016. Les participants à l'enquête alimentaire n'étaient pas différents de ceux de la cohorte GrippeNet.fr. L'inclusion et la collecte des données se sont déroulées en une semaine entre le 22/02/2016 et le 26/02/2016, sur les aliments consommés en période de fêtes et en période hivernale classique. Les participants qui avaient déclaré un voyage à l'étranger et ceux qui présentaient des symptômes gastro-intestinaux étaient exclus des analyses. Les analyses étaient ajustées sur l'âge et le sexe car la répartition des témoins différait par rapport aux cas selon ces deux caractéristiques. La comparaison des consommations alimentaires des participants de GrippeNet.fr a permis de démontrer que la fréquence de consommation de deux fromages au lait cru était 4 à 5 fois supérieure chez les personnes atteintes de salmonellose comparée aux GrippeNauts qui avaient participé à cette enquête. Les résultats obtenus grâce à cette étude ont conforté les premières hypothèses de SpF. Elles ont permis la mise en place de mesures de contrôle dans les lieux de fabrication de ces fromages par la Direction générale de l'alimentation (DGAL) et de ses services de contrôle.

6.2. GrippeNet.fr au-delà des maladies infectieuses

Au-delà du suivi de pathologies définies, les systèmes participatifs sont des outils efficaces pour sensibiliser la population à la santé publique [129], et promouvoir la santé et l'éducation à la santé de façon ciblée ou globale [136, 137]. En effet, en dehors de la grippe et des maladies infectieuses, GrippeNet.fr s'est avérée utile pour l'évaluation de mesures de prévention et de communication. Plusieurs projets sont actuellement en cours de réalisation.

Un premier travail porte sur **les connaissances et comportements face à la grippe**. Cette étude a été conduite lors des saisons 2016/2017 et 2017/2018, dans trois pays européens (France, Italie, Royaume-Uni). L'objectif était d'étudier l'association entre la perception de risque et l'anxiété ressentie face à la grippe, et d'évaluer le changement de comportements lors de l'épidémie saisonnière. Pour cette étude, des questions ont été ajoutées au questionnaire d'inclusion, et trois questionnaires supplémentaires ont été envoyés aux participants durant la saison : lors de la phase ascendante, avant le franchissement du seuil épidémique selon les indications du réseau Sentinelles ; lors de la phase ascendante juste avant l'atteinte du pic épidémique ; et en phase descendante de l'épidémie. Les données collectées étaient les variables sociodémographiques et

des variables liées à : la connaissance (i.e. niveau de connaissance et sources primaires d'information sur la grippe) ; l'anxiété (i.e. impact concret au quotidien et ressenti) ; l'efficacité perçue des mesures de prévention (i.e. niveau de confiance dans les mesures barrières de type hygiène, vaccination). Sur les 5 517 participants éligibles de la saison 2016/2017, 48% (n=2661) ont rempli les trois questionnaires supplémentaires. Plus de la majorité des participants pensaient être plus à risque d'attraper la grippe, mais avaient confiance dans toutes les mesures préventives. L'anxiété impactait leur vie quotidienne dans plus d'un cas sur deux. Certains ont changé leur comportement au fil de la saison. Autour du pic épidémique, près de la moitié ont essayé d'éviter les personnes grippées. Ils ont mis en pratique certaines mesures d'hygiène comme l'utilisation de mouchoirs pour tousser ou éternuer, ou l'utilisation de solutions hydro-alcooliques. Par contre, les participants n'ont pas réduit l'usage des transports en commun ou l'annulation d'événements sociaux. Les raisons de ces changements de comportement étaient liées aux recommandations faites par les professionnels de santé et à la présence de symptômes grippaux. Ces analyses préliminaires ont pour but d'être complétées pour connaître les facteurs associés au changement de comportement à l'approche du pic. Cette étude a pour but de mieux axer les campagnes de prévention, promotion et communication de santé publique

Le second porte sur **le port du masque chez les personnes qui présentent un syndrome grippal**. Cette étude a été mise en place conjointement par SpF et l'équipe en charge de GrippeNet.fr lors de la dernière saison 2018/2019. Son objectif est d'évaluer l'acceptabilité et l'adhésion au port du masque comme mesure barrière chez les sujets présentant une maladie grippale en population générale. Les facteurs associés à l'acceptabilité et à l'adhésion du port de masque seront étudiés en population générale, avec un focus chez les enfants, connus pour être des vecteurs de transmission des pathologies grippales, ainsi que les freins liés au port du masque.

D'autres études sont également menées afin de recueillir la perception, l'opinion de la population sur des thématiques de santé variées. Un travail a ainsi été réalisé pour étudier la **perception de la population sur l'extension de l'obligation vaccinale** mise en place en France en janvier 2018. Connaissant le contexte français par rapport à l'hésitation vaccinale, il paraissait important pour les autorités de santé de mesurer l'impact de cette politique de santé publique avant son instauration. L'étude a donc été conduite entre novembre et décembre 2017 parmi les GrippeNautes de 18 ans et plus. Sur les 6 383 participants contactés, 3222 ont complété le questionnaire dédié, ce qui représente un taux de participation de 50,5%. Après ajustement sur l'âge, le sexe et le niveau d'éducation, 64,5% de la population française était en faveur de l'élargissement de l'obligation vaccinale [207]. Une étude menée en 2008 auprès de la population générale en France avait montré que 56% de la population était en faveur de cette politique vaccinale [208]. La perception de l'innocuité et des avantages des vaccins étaient des prédicteurs majeurs d'une opinion positive de la population à l'égard de cette nouvelle politique.

D'autres travaux destinés à recueillir l'opinion de la population sur des thématiques de santé différentes sont en cours de réalisation, comme la surmédicalisation ou la perception des risques d'addiction aux opioïdes.

Les messages clés du Chapitre 6

- La surveillance participative via GrippeNet.fr permet le suivi et l'analyse d'autres pathologies infectieuses que celles pour lesquelles le système a été mis en place.
- La flexibilité et la rapidité du système font de lui un outil d'investigation efficace et précoce.
- GrippeNet.fr s'est avérée utile pour l'évaluation de mesures de prévention et de communication mises en place par les institutions sanitaires.

Conclusions

Les travaux de recherche réalisés au cours de cette thèse et ceux menés en parallèle par le biais de la cohorte participative GrippeNet.fr visent à évaluer et valider cette nouvelle source de données, en mettant en évidence ses bénéfices dans le champ de la surveillance en santé publique, en complément des surveillances existantes. L'intérêt de ce système participatif a été étudié à un niveau populationnel, avec l'analyse des incidences des syndromes grippaux aux niveaux national et régional, et à niveau individuel, avec l'étude des caractéristiques des individus avec un syndrome grippal et de la démographie des personnes affectées.

Atouts de la surveillance participative : réactivité et précision

Les travaux menés grâce à cette thèse ont permis de mettre en évidence la fiabilité de la cohorte GrippeNet.fr dans le champ de la surveillance sanitaire en population générale. Pour valider ce nouveau système et les signaux captés, une comparaison avec les systèmes traditionnels de référence a été nécessaire, bien que chaque système ait des caractéristiques propres.

Le premier travail réalisé a conduit à mieux comprendre le signal des syndromes grippaux émis par la cohorte GrippeNet.fr en prenant en compte des critères liés aux syndromes grippaux eux-mêmes (i.e. variation des différentes définitions de cas), et des critères de calcul des incidences propres à cette nouvelle source de données participative (i.e. critères liés au rapport des symptômes, critères liés à la participation). Cette étude de sensibilité menée sur les différents critères qui impactent le calcul des incidences a permis de mettre en évidence un cadre d'analyse standardisé de calcul des incidences directement à partir de la population. Si les analyses ont été réalisées de façon rétrospective, un travail visant à mettre en application la méthode de calcul élaborée à partir de données collectées en temps réel est prévu. En effet, il est important de pouvoir évaluer l'efficacité du système de façon prospective pour avoir une évaluation globale de cette nouvelle surveillance [65]. La comparaison avec les systèmes traditionnels a été réalisée à partir de deux caractéristiques : la tendance et l'amplitude du signal. La comparaison des tendances s'est faite par l'étude de la corrélation avec les systèmes traditionnels en France. Ceci a permis de montrer que GrippeNet.fr pouvait déjà suivre l'évolution des syndromes grippaux, quelle que soit la définition utilisée grâce à la flexibilité des données collectées, mais aussi a mis en évidence la rapidité du système à estimer les syndromes grippaux dans la population pour une semaine donnée. Le pic épidémique était souvent anticipé d'une semaine. Ces facultés ont été retrouvées dans d'autres systèmes participatifs [138, 150, 152, 153]. La seconde caractéristique du système concerne l'amplitude de l'épidémie suivie par rapport à celle des systèmes traditionnels. Comme attendue celle-ci était plus grande avec GrippeNet.fr, puisque l'on intègre ici la population générale et pas seulement la population qui consulte un professionnel de santé

pour un syndrome grippal. L'analyse réalisée sur deux saisons n'a pas permis de donner des informations sur l'ampleur des épidémies suivies en population comme aucune référence n'était disponible et que cela dépend des hypothèses avancées sur les différents critères (notamment la participation). Ici, l'objectif premier était de visualiser et comprendre le signal des syndromes grippaux perçus. Avec huit saisons de recul, il est désormais possible de mettre en place le calcul d'un seuil ou de niveaux d'activité spécifiques à la population, pour être en mesure de juger l'activité grippale au regard de l'historique des épidémies suivies via GrippeNet.fr.

Le second travail s'est focalisé sur la compréhension des données individuelles, grâce à la collecte de données épidémiologiques fines, telles que les caractéristiques sociodémographiques, environnementales, ou de santé de cette cohorte participative en population. L'objectif était de valider les données épidémiologiques collectées via ce système, en déterminant les facteurs associés aux syndromes grippaux en population générale, et en les comparant à ceux retrouvés dans la littérature. Ces derniers sont habituellement définis par le caractère sévère du syndrome grippal, comme les études sont principalement menées à partir du système de soins. La connaissance des facteurs de risque d'infection avec un degré de sévérité moindre, dans une population plus large et en temps réel représente un atout important pour les autorités sanitaires qui peuvent élaborer des réponses adaptées et rapides. La plupart des facteurs retrouvés par notre travail étaient similaires à ceux de la littérature, montrant que certains facteurs sont déjà des facteurs d'infection avant d'être des facteurs de sévérité. De plus, la finesse des données collectées nous a permis d'en décrire des nouveaux et d'autres plus débattus, que nous pourrions étudier plus spécifiquement pour mieux les comprendre. Cette étude a finalement permis de valider un peu plus les données collectées via la cohorte participative GrippeNet.fr.

Les deux travaux réalisés ont montré que GrippeNet.fr pouvait suivre l'évolution des syndromes grippaux lors des épidémies saisonnières de façon efficace et efficiente, à l'échelle populationnelle comme à l'échelle individuelle. Les systèmes de surveillance traditionnels de la grippe continueront à représenter le « gold-standard » en matière de surveillance grâce à leur spécificité, mais les systèmes participatifs ou issus d'Internet et des réseaux sociaux ont un rôle à jouer en complément. Si le RSI et le programme de sécurité sanitaire mondiale promeuvent que les pays se dotent de la capacité de pouvoir signaler rapidement des événements infectieux émergents, l'inclusion des méthodes de surveillance participative dans les pays qui cherchent à atteindre ses objectifs paraît une bonne alternative à la surveillance faite à partir d'Internet et des réseaux sociaux. En effet, les systèmes participatifs permettent de contrôler les données collectées puisque les questionnaires sont élaborés par des professionnels de santé dans le but d'en effectuer un suivi épidémiologique. L'utilisation d'Internet n'est ici qu'un moyen de pouvoir rendre les données et résultats accessibles rapidement, alors que l'analyse des données collectées s'effectue comme pour le suivi de cohortes épidémiologiques traditionnelles. Ces propriétés ont pu être testées au-delà de la veille sanitaire et de la recherche sur les maladies grippales. La cohorte GrippeNet.fr s'est avérée utile pour suivre et étudier d'autres pathologies infectieuses grâce aux divers symptômes collectés couplés aux données épidémiologiques recueillies. Elle permet également de mesurer les comportements et perceptions directement à

partir de la population générale, ce qui est fondamental pour la mise en place d'études pilotes évaluant l'acceptation de certaines actions de santé publique avant leur diffusion à large échelle. Par exemple, l'étude sur l'acceptation du port du masque a été menée d'abord au sein de la cohorte GrippeNet.fr car les autorités sanitaires voulaient évaluer son intérêt avant de mener une campagne de communication nationale.

GrippeNet.fr et les cohortes participatives via Internet représentent de bons outils pour la surveillance, le suivi épidémiologique, et la perception de la population sur des thématiques variées, avec d'un côté la rigueur des systèmes traditionnels, et de l'autre la souplesse et la rapidité de recueil des données issues d'Internet. Toutes les sources de données comprennent des biais et des limites. L'essentiel est de bien les connaître et de les comprendre avant de les utiliser dans le cadre de la surveillance sanitaire [137]. A travers ces deux travaux, et en complément de celui effectué sur la représentativité des participants de GrippeNet.fr, nous avons voulu valider ce système de surveillance participatif pour le suivi des syndromes grippaux, et pour son utilisation dans le système de surveillance de la grippe en France. GrippeNet.fr permet de combler le dernier niveau de surveillance pour lequel aucune surveillance n'était établie. Il est en ce sens un bon candidat pour le suivi des syndromes grippaux directement auprès de la population.

Un rôle à jouer dans la surveillance multi-sources et la modélisation

Les études visant à évaluer la qualité, la réactivité ou la représentativité de l'information issue des systèmes participatifs doivent promouvoir l'adoption et l'intégration de ces systèmes au suivi systématique de la santé dans la communauté [136]. Si cela reste encore un défi, des travaux ont été menés récemment pour mettre en évidence l'intérêt d'une surveillance multi-sources pour mieux suivre et prédire l'évolution des syndromes grippaux [87, 209]. Un système de surveillance multi-sources (données traditionnelles, surveillance syndromique et digitale) basé sur l'apprentissage automatique (i.e. machine learning) permettrait le meilleur des suivis (Figure 13) [87, 209-211]. L'idée est effectivement de mettre à profit l'utilisation conjointe de données de sources différentes pour améliorer l'estimation des incidences des syndromes grippaux et des prévisions, sans pour autant substituer une source de données par une autre. En France, le système MASS mis en place par SpF est basé sur l'intégration de trois sources de données issues du système de soins. Il permet de suivre le niveau d'activité des syndromes grippaux au cours de la saison hivernale, en synthétisant les informations collectées à partir des trois sources de données [32]. En parallèle, le réseau Sentinelles utilise les données de délivrance de médicaments pour prédire les incidences des syndromes grippaux [212]. L'ajout des données issues de GrippeNet.fr à ces différents projets d'analyse multi-sources permettrait d'étudier le réel bénéfice de cette surveillance dans le champ de la surveillance sanitaire, et de justifier son utilisation par les autorités de santé. Dans ce contexte, de récents projets de recherche en modélisation à partir de données multi-sources ont été réalisés [117, 136, 209, 213]. Ces projets ont été menés pour comparer d'un côté la valeur ajoutée d'une source par rapport à l'autre [117],

et de l'autre pouvoir comparer la combinaison optimale de sources de données pour la prévision en temps réel [209, 213]. *Wagner et al.* montraient que les données issues de Google permettaient un suivi plus précis des syndromes grippaux que les données issues de Twitter [117]. *Reich et al.* ont mis en évidence que la majorité des modèles multi-sources étudiés pouvait estimer des prévisions en temps réel de façon plus précise, même si les prévisions à court-terme (une ou deux semaines) restaient plus fiables que celles effectuées à long-terme (trois ou quatre semaines) [209]. Cependant, à notre connaissance, une seule étude a intégré différentes sources de données dont une issue d'un système participatif (Flu Near You) dans le cadre d'une surveillance multi-sources de la grippe [210]. La combinaison de données de surveillance participative avec d'autres sources de données, couplée avec une approche de modélisation par apprentissage automatique, permettait d'améliorer considérablement les prévisions à court terme de l'activité grippale [210]. Des travaux similaires pourraient être adaptés en France, avec l'utilisation des données de GrippeNet.fr.

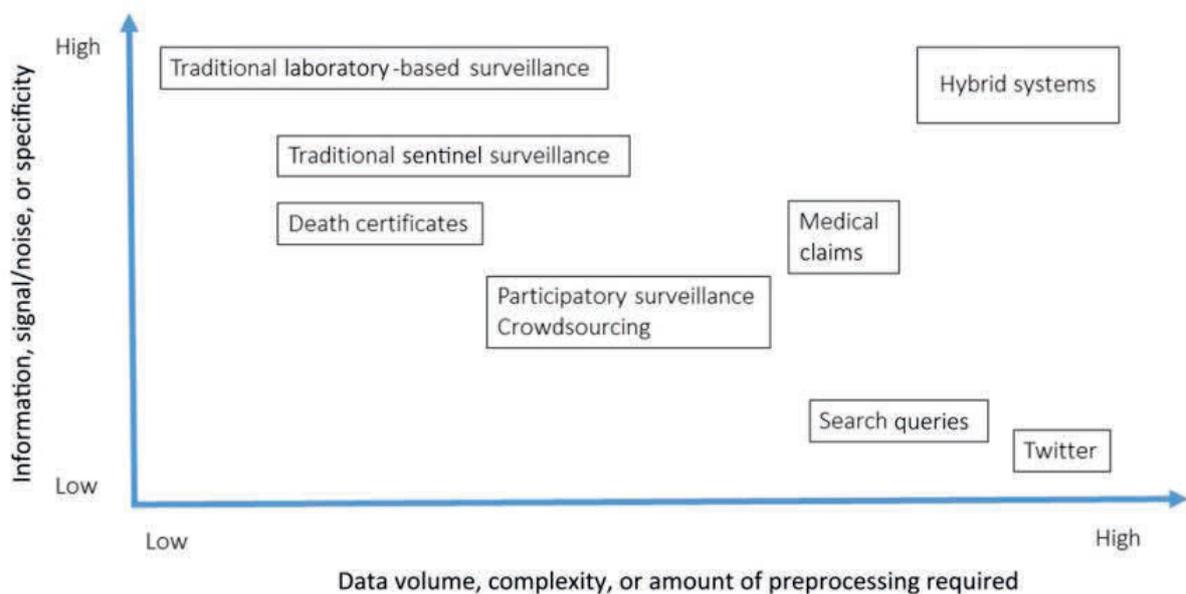


Figure 13. Aperçu des caractéristiques des systèmes de surveillance des maladies infectieuses. (Les systèmes hybrides combinant la surveillance traditionnelle et les données massives issues d'Internet représentent le "meilleur" des systèmes, avec un retour d'information et un volume de données élevés). Sources: *Simonsen et al.* [89]

Les études menées jusqu'à présent avec les données issues du système GrippeNet.fr ont permis de montrer la valeur ajoutée de ce système de surveillance sanitaire en population. Même si des études complémentaires doivent être réalisées, comme par exemple l'analyse prospective des données, il est important de réfléchir désormais à l'intégration de cette nouvelle source dans le dispositif de surveillance de la grippe existant, afin d'avoir un système complet, réactif et efficient.

Références

1. Astagneau P, Ancelle T. *Surveillance épidémiologique : principes, méthodes et applications en santé publique*. 2011.
2. Langmuir AD. *The surveillance of communicable diseases of national importance*. N Engl J Med, 1963. **268**: p. 182-92.
3. Henderson DA. *The Development of Surveillance Systems*. Am J Epidemiol, 2016. **183**(5): p. 381-6.
4. Park HA, et al. *Digital Epidemiology: Use of Digital Data Collected for Non-epidemiological Purposes in Epidemiological Studies*. Healthc Inform Res, 2018. **24**(4): p. 253-262.
5. Desenclos JC. *[Surveillance of infectious diseases: principles and organisation in France in 2005]*. Med Mal Infect, 2005. **35**(5): p. 232-44.
6. Butler D. *When Google got flu wrong*. Nature, 2013. **494**(7436): p. 155-6.
7. Olson DR, et al. *Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales*. PLoS Comput Biol, 2013. **9**(10): p. e1003256.
8. Choi BC. *The past, present, and future of public health surveillance*. Scientifica (Cairo), 2012. **2012**: p. 875253.
9. Henning KJ. *What is syndromic surveillance?* MMWR Suppl, 2004. **53**: p. 5-11.
10. World Health Organisation. *Early detection, assessment and response to acute public health events: implementation of Early Warning and Response with a focus on Event-Based Surveillance*. 2014, WHO: Geneva.
11. Baumgartner A, et al. *Seasonality, timing, and climate drivers of influenza activity worldwide*. J Infect Dis, 2012. **206**(6): p. 838-46.
12. Organisation mondiale de la Santé. *Grippe saisonnière*. Accessed : 25/01/2019; Available from: [https://www.who.int/fr/news-room/fact-sheets/detail/influenza-\(seasonal\)](https://www.who.int/fr/news-room/fact-sheets/detail/influenza-(seasonal)).
13. Iuliano AD, et al. *Estimates of global seasonal influenza-associated respiratory mortality: a modelling study*. Lancet, 2018. **391**(10127): p. 1285-1300.
14. Souty C, et al. *Influenza epidemics observed in primary care from 1984 to 2017 in France: A decrease in epidemic size over time*. Influenza Other Respir Viruses, 2018.
15. Szucs T. *The socio-economic burden of influenza*. J Antimicrob Chemother, 1999. **44 Suppl B**: p. 11-5.
16. Carrat F, et al. *Influenza burden of illness: estimates from a national prospective survey of household contacts in France*. Arch Intern Med, 2002. **162**(16): p. 1842-8.
17. Gessner BD, Shindo N, and Briand S. *Seasonal influenza epidemiology in sub-Saharan Africa: a systematic review*. Lancet Infect Dis, 2011. **11**(3): p. 223-35.
18. Institut de veille sanitaire. *La veille et l'alerte sanitaires en France*. 2011. p. 60.
19. Capek I. *Orientations pour la révision des modalités de surveillance des maladies transmissibles en France*. Bull Epidémiol Hebd, 25 juin 1996. **26**.
20. Seringe E, Astagneau P. *Structure et organisation du système de surveillance en France : comparaison avec les pays anglo-saxons*, in *Surveillance épidémiologique : principes, méthodes et applications en santé publique*, M.S. Publications, Editor. 2011.
21. Aguilera JF, et al. *Heterogeneous case definitions used for the surveillance of influenza in Europe*. Eur J Epidemiol, 2003. **18**(8): p. 751-4.
22. Jiang L, et al. *Performance of case definitions for influenza surveillance*. Euro Surveill, 2015. **20**(22): p. 21145.

23. Casalegno JS, et al. *Performance of influenza case definitions for influenza community surveillance: based on the French influenza surveillance network GROG, 2009-2014*. Euro Surveill, 2017. **22**(14).
24. Souty C, et al. *Baseline characteristics and clinical symptoms related to respiratory viruses identified among patients presenting with influenza-like illness in primary care*. Clin Microbiol Infect, 2019.
25. Bamberger E, et al. *What is the clinical relevance of respiratory syncytial virus bronchiolitis?: findings from a multi-center, prospective study*. Eur J Clin Microbiol Infect Dis, 2012. **31**(12): p. 3323-30.
26. Falsey AR, et al. *Respiratory syncytial virus infection in elderly and high-risk adults*. N Engl J Med, 2005. **352**(17): p. 1749-59.
27. Hay AJ, McCauley JW. *The WHO global influenza surveillance and response system (GISRS)-A future perspective*. Influenza Other Respir Viruses, 2018.
28. Organisation mondiale de la Santé. *Règlement sanitaire international (2005)*. 3rd ed. 2016. Accessed : 25/01/2019; Available from: <http://apps.who.int/iris/bitstream/10665/246187/1/9789242580495-fre.pdf?ua=1>.
29. Declich S, Carter AO. *Public health surveillance: historical origins, methods and evaluation*. Bull World Health Organ, 1994. **72**(2): p. 285-304.
30. Organisation mondiale de la Santé. *Influenza Surveillance and lab network*. 2018. Accessed: 28/01/2019; Available from: <http://www.euro.who.int/en/health-topics/communicable-diseases/influenza/surveillance-and-lab-network>.
31. Nielsen J, Krause TG, Molbak K. *Influenza-associated mortality determined from all-cause mortality, Denmark 2010/11-2016/17: The FluMOMO model*. Influenza Other Respir Viruses, 2018.
32. Pelat C, et al. *Improving regional influenza surveillance through a combination of automated outbreak detection methods: the 2015/16 season in France*. Euro Surveill, 2017. **22**(32).
33. Caserio-Schönemann C, et al. *Le système de surveillance syndromique SurSaUD*. Bull Epidémiol Hebd, 2014. **3-4**: p. 38-44.
34. Deckers JG, et al. *European primary care surveillance networks: their structure and operation*. Fam Pract, 2006. **23**(2): p. 151-8.
35. Mosnier A, et al. *Ten influenza seasons in France: distribution and timing of influenza A and B circulation, 2003-2013*. BMC Infect Dis, 2015. **15**: p. 357.
36. Flahault A, et al. *Virtual surveillance of communicable diseases: a 20-year experience in France*. Stat Methods Med Res, 2006. **15**(5): p. 413-21.
37. Valleron AJ, et al. *A computer network for the surveillance of communicable diseases: the French experiment*. Am J Public Health, 1986. **76**(11): p. 1289-92.
38. Réseau Sentinelles. *Bilan annuel 2018-2019*. Available from: <https://www.sentiweb.fr/document/4633>
39. Turbelin C, Boelle PY. *Improving general practice based epidemiologic surveillance using desktop clients: the French Sentinel Network experience*. Stud Health Technol Inform, 2010. **160**(Pt 1): p. 442-6.
40. Riviere M, et al. *Secular trends in incidence of acute gastroenteritis in general practice, France, 1991 to 2015*. Euro Surveill, 2017. **22**(50).
41. Costagliola D, et al. *A routine tool for detection and assessment of epidemics of influenza-like syndromes in France*. Am J Public Health, 1991. **81**(1): p. 97-9.
42. Pelat C, et al. *Online detection and quantification of epidemics*. BMC Med Inform Decis Mak, 2007. **7**: p. 29.
43. Viboud C, et al. *Prediction of the spread of influenza epidemics by the method of analogues*. Am J Epidemiol, 2003. **158**(10): p. 996-1006.
44. Gonzalez Chiappe S, et al. *Herpes zoster: Burden of disease in France*. Vaccine, 2010. **28**(50): p. 7933-8.
45. Souty C, et al. *Vaccination against varicella as post-exposure prophylaxis in adults: a quantitative assessment*. Vaccine, 2015. **33**(3): p. 446-50.

46. Mandl KD, et al. *Implementing syndromic surveillance: a practical guide informed by the early experience.* J Am Med Inform Assoc, 2004. **11**(2): p. 141-50.
47. Magruder SF, et al. *Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance.* MMWR Suppl, 2004. **53**: p. 117-22.
48. Hughes HE, et al. *Syndromic Surveillance Revolution? Public Health Benefits of Modernizing the Emergency Care Patient Health Record in England.* Public Health Rep, 2017. **132**(1_suppl): p. 12s-15s.
49. Simonsen L, et al. *The impact of influenza epidemics on hospitalizations.* J Infect Dis, 2000. **181**(3): p. 831-7.
50. Lazarus R, et al. *Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): the example of lower respiratory infection.* BMC Public Health, 2001. **1**: p. 9.
51. Viboud C, et al. *Transmissibility and mortality impact of epidemic and pandemic influenza, with emphasis on the unusually deadly 1951 epidemic.* Vaccine, 2006. **24**(44-46): p. 6701-7.
52. *Assessment of syndromic surveillance in Europe.* Lancet, 2011. **378**(9806): p. 1833-4.
53. Lewis MD, et al. *Disease outbreak detection system using syndromic data in the greater Washington DC area.* Am J Prev Med, 2002. **23**(3): p. 180-6.
54. Retel O, et al. *Contribution des associations SOS Médecins à une surveillance locale de la grippe saisonnière en France.* Bull Epidémiol Hebd, 2014. **28**: p. 466-72.
55. Gault G, et al. *Performance of a syndromic system for influenza based on the activity of general practitioners, France.* J Public Health (Oxf), 2009. **31**(2): p. 286-92.
56. Souty C, Boelle PY. *Improving incidence estimation in practice-based sentinel surveillance networks using spatial variation in general practitioner density.* BMC Med Res Methodol, 2016. **16**(1): p. 156.
57. Starfield B. *Is primary care essential?* Lancet, 1994. **344**(8930): p. 1129-33.
58. Wagner MM, et al. *Design of a national retail data monitor for public health surveillance.* J Am Med Inform Assoc, 2003. **10**(5): p. 409-18.
59. Hope K, et al. *Syndromic Surveillance: is it a useful tool for local outbreak detection?* J Epidemiol Community Health, 2006. **60**(5): p. 374-5.
60. Lipsitch M, et al. *How to maintain surveillance for novel influenza A H1N1 when there are too many cases to count.* Lancet, 2009. **374**(9696): p. 1209-11.
61. Smith GE, et al. *Syndromic surveillance: two decades experience of sustainable systems - its people not just data!* Epidemiol Infect, 2019. **147**: p. e101.
62. Caserio-Schonemann C, Meynard JB. *Ten years experience of syndromic surveillance for civil and military public health, France, 2004-2014.* Euro Surveill, 2015. **20**(19): p. 35-8.
63. Yoon PW, Ising AI, Gunn JE. *Using Syndromic Surveillance for All-Hazards Public Health Surveillance: Successes, Challenges, and the Future.* Public Health Rep, 2017. **132**(1_suppl): p. 3s-6s.
64. Wagner MM, et al. *The emerging science of very early detection of disease outbreaks.* J Public Health Manag Pract, 2001. **7**(6): p. 51-9.
65. Milinovich GJ, et al. *Internet-based surveillance systems for monitoring emerging infectious diseases.* Lancet Infect Dis, 2014. **14**(2): p. 160-8.
66. Ekman A, Litton JE. *New times, new needs; e-epidemiology.* Eur J Epidemiol, 2007. **22**(5): p. 285-92.
67. Brownstein JS, Freifeld CC, Madoff LC. *Digital disease detection--harnessing the Web for public health surveillance.* N Engl J Med, 2009. **360**(21): p. 2153-5, 2157.
68. Salathe M. *Digital epidemiology: what is it, and where is it going?* Life Sci Soc Policy, 2018. **14**(1): p. 1.
69. Samerski S. *Individuals on alert: digital epidemiology and the individualization of surveillance.* Life Sci Soc Policy, 2018. **14**(1): p. 13.

70. Eysenbach G. *Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet*. J Med Internet Res, 2009. **11**(1): p. e11.
71. Barboza P, et al. *Factors influencing performance of internet-based biosurveillance systems used in epidemic intelligence for early detection of infectious diseases outbreaks*. PLoS One, 2014. **9**(3): p. e90536.
72. Gajewski KN, et al. *A review of evaluations of electronic event-based biosurveillance systems*. PLoS One, 2014. **9**(10): p. e111222.
73. Heymann DL, Rodier GR. *Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases*. Lancet Infect Dis, 2001. **1**(5): p. 345-53.
74. Keller M, et al. *Use of unstructured event-based reports for global infectious disease surveillance*. Emerg Infect Dis, 2009. **15**(5): p. 689-95.
75. Madoff LC. *ProMED-mail: an early warning system for emerging diseases*. Clin Infect Dis, 2004. **39**(2): p. 227-32.
76. Carrion M, Madoff LC. *ProMED-mail: 22 years of digital surveillance of emerging infectious diseases*. Int Health, 2017. **9**(3): p. 177-183.
77. Milne-Price S, Miazgowicz KL, Munster VJ. *The emergence of the Middle East respiratory syndrome coronavirus*. Pathog Dis, 2014. **71**(2): p. 121-36.
78. Mykhalovskiy E, Weir L. *The Global Public Health Intelligence Network and early warning outbreak detection: a Canadian contribution to global public health*. Can J Public Health, 2006. **97**(1): p. 42-4.
79. Dion M, AbdelMalik P, Mawudeku A. *Big Data and the Global Public Health Intelligence Network (GPHIN)*. Can Commun Dis Rep, 2015. **41**(9): p. 209-214.
80. Eysenbach G. *SARS and population health technology*. J Med Internet Res, 2003. **5**(2): p. e14.
81. Brownstein JS, Freifeld CC. *HealthMap: the development of automated real-time internet surveillance for epidemic intelligence*. Euro Surveill, 2007. **12**(11): p. E071129.5.
82. Freifeld CC, et al. *HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports*. J Am Med Inform Assoc, 2008. **15**(2): p. 150-7.
83. Brownstein JS, et al. *Surveillance Sans Frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project*. PLoS Med, 2008. **5**(7): p. e151.
84. Nelson NP, Brownstein JS, Hartley DM. *Event-based biosurveillance of respiratory disease in Mexico, 2007-2009: connection to the 2009 influenza A(H1N1) pandemic?* Euro Surveill, 2010. **15**(30).
85. Hossain N, Househ M. *Using HealthMap to Analyse Middle East Respiratory Syndrome (MERS) Data*. Stud Health Technol Inform, 2016. **226**: p. 213-6.
86. McGough SF, et al. *Forecasting Zika Incidence in the 2016 Latin America Outbreak Combining Traditional Disease Surveillance with Search, Social Media, and News Report Data*. PLoS Negl Trop Dis, 2017. **11**(1): p. e0005295.
87. Brownstein JS, et al. *Combining Participatory Influenza Surveillance with Modeling and Forecasting: Three Alternative Approaches*. JMIR Public Health Surveill, 2017. **3**(4): p. e83.
88. Huston JE, et al. *Searching the Web for Influenza Vaccines: HealthMap Vaccine Finder*. Am J Public Health, 2015. **105**(8): p. e134-9.
89. Simonsen L, et al. *Infectious Disease Surveillance in the Big Data Era: Towards Faster and Locally Relevant Systems*. J Infect Dis, 2016. **214**(suppl_4): p. S380-s385.
90. Polgreen PM, et al. *Using internet searches for influenza surveillance*. Clin Infect Dis, 2008. **47**(11): p. 1443-8.
91. Ginsberg J, et al. *Detecting influenza epidemics using search engine query data*. Nature, 2009. **457**(7232): p. 1012-4.

92. Cook S, et al. *Assessing Google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic*. PLoS One, 2011. **6**(8): p. e23610.
93. Lazer D, et al. *Big data. The parable of Google Flu: traps in big data analysis*. Science, 2014. **343**(6176): p. 1203-5.
94. Lampos V, et al. *Advances in nowcasting influenza-like illness rates using search query logs*. Sci Rep, 2015. **5**: p. 12760.
95. Yuan Q, et al. *Monitoring influenza epidemics in china with search query from baidu*. PLoS One, 2013. **8**(5): p. e64323.
96. Generous N, et al. *Global disease monitoring and forecasting with Wikipedia*. PLoS Comput Biol, 2014. **10**(11): p. e1003892.
97. Sharpe JD, et al. *Evaluating Google, Twitter, and Wikipedia as Tools for Influenza Surveillance Using Bayesian Change Point Analysis: A Comparative Analysis*. JMIR Public Health Surveill, 2016. **2**(2): p. e161.
98. McIver DJ, Brownstein JS. *Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time*. PLoS Comput Biol, 2014. **10**(4): p. e1003581.
99. Pelat C, et al. *More diseases tracked by using Google Trends*. Emerg Infect Dis, 2009. **15**(8): p. 1327-8.
100. Althouse BM, Ng YY, Cummings DA. *Prediction of dengue incidence using search query surveillance*. PLoS Negl Trop Dis, 2011. **5**(8): p. e1258.
101. Chan EH, et al. *Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance*. PLoS Negl Trop Dis, 2011. **5**(5): p. e1206.
102. Ocampo AJ, Chunara R, Brownstein JS. *Using search queries for malaria surveillance, Thailand*. Malar J, 2013. **12**: p. 390.
103. Seifter A, et al. *The utility of "Google Trends" for epidemiological research: Lyme disease as an example*. Geospat Health, 2010. **4**(2): p. 135-7.
104. Kapitany-Foveny M, et al. *Can Google Trends data improve forecasting of Lyme disease incidence?* Zoonoses Public Health, 2019. **66**(1): p. 101-107.
105. McCarthy MJ. *Internet monitoring of suicide risk in the population*. J Affect Disord, 2010. **122**(3): p. 277-9.
106. Foroughi F, et al. *"Googling" for Cancer: An Infodemiological Assessment of Online Search Interests in Australia, Canada, New Zealand, the United Kingdom, and the United States*. JMIR Cancer, 2016. **2**(1): p. e5.
107. McLean S, Lennon P, Glare P. *Internet search query analysis can be used to demonstrate the rapidly increasing public awareness of palliative care in the USA*. BMJ Support Palliat Care, 2019. **9**(1): p. 40-44.
108. Grajales FJ, et al. *Social media: a review and tutorial of applications in medicine and health care*. J Med Internet Res, 2014. **16**(2): p. e13.
109. Corley CD, et al. *Text and structural data mining of influenza mentions in Web and social media*. Int J Environ Res Public Health, 2010. **7**(2): p. 596-615.
110. Sinnenberg L, et al. *Twitter as a Tool for Health Research: A Systematic Review*. Am J Public Health, 2017. **107**(1): p. e1-e8.
111. Collier N, Son NT, Nguyen NM. *OMG U got flu? Analysis of shared health messages for bio-surveillance*. J Biomed Semantics, 2011. **2** Suppl 5: p. S9.
112. Chew C, Eysenbach G. *Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak*. PLoS One, 2010. **5**(11): p. e14118.
113. Signorini A, Segre AM, Polgreen PM. *The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic*. PLoS One, 2011. **6**(5): p. e19467.

114. Broniatowski DA, Paul MJ, Dredze M. *National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic*. PLoS One, 2013. **8**(12): p. e83672.
115. Aslam AA, et al. *The reliability of tweets as a supplementary method of seasonal influenza surveillance*. J Med Internet Res, 2014. **16**(11): p. e250.
116. Nagar R, et al. *A case study of the New York City 2012-2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives*. J Med Internet Res, 2014. **16**(10): p. e236.
117. Wagner M, et al. *The added value of online user-generated content in traditional methods for influenza surveillance*. Sci Rep, 2018. **8**(1): p. 13963.
118. Salathe M, et al. *Digital epidemiology*. PLoS Comput Biol, 2012. **8**(7): p. e1002616.
119. Althouse BM, et al. *Enhancing disease surveillance with novel data streams: challenges and opportunities*. EPJ Data Sci, 2015. **4**.
120. Charles-Smith LE, et al. *Using Social Media for Actionable Disease Surveillance and Outbreak Management: A Systematic Literature Review*. PLoS One, 2015. **10**(10): p. e0139701.
121. Davidson MW, Haim DA, Radin JM. *Using networks to combine "big data" and traditional surveillance to improve influenza predictions*. Sci Rep, 2015. **5**: p. 8154.
122. Chabot-Couture G, et al. *Advancing digital methods in the fight against communicable diseases*. Int Health, 2015. **7**(2): p. 79-81.
123. Hulth A, Rydevik G, Linde A. *Web queries as a source for syndromic surveillance*. PLoS One, 2009. **4**(2): p. e4378.
124. Elliot AJ, et al. *Internet-based remote health self-checker symptom data as an adjuvant to a national syndromic surveillance system*. Epidemiol Infect, 2015. **143**(16): p. 3416-22.
125. Keramarou M, et al. *Two waves of pandemic influenza A(H1N1) 2009 in Wales--the possible impact of media coverage on consultation rates, April-December 2009*. Euro Surveill, 2011. **16**(3).
126. Moorhead SA, et al. *A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication*. J Med Internet Res, 2013. **15**(4): p. e85.
127. Swan M. *Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem*. J Med Internet Res, 2012. **14**(2): p. e46.
128. Behrend TS, et al. *The viability of crowdsourcing for survey research*. Behav Res Methods, 2011. **43**(3): p. 800-13.
129. Wazny K. *Applications of crowdsourcing in health: an overview*. J Glob Health, 2018. **8**(1): p. 010502.
130. Boulos K, et al. *Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples*. Int J Health Geogr, 2011. **10**: p. 67.
131. Ponsonby AL, Mattingly K. *Evaluating New Ways of Working Collectively in Science with a Focus on Crowdsourcing*. EBioMedicine, 2015. **2**(7): p. 627-8.
132. Nguyen VT, et al. *A scoping review provided a framework for new ways of doing research through mobilizing collective intelligence*. J Clin Epidemiol, 2019.
133. Crequit P, et al. *Mapping of Crowdsourcing in Health: Systematic Review*. J Med Internet Res, 2018. **20**(5): p. e187.
134. Leal-Neto OB, et al. *Digital disease detection and participatory surveillance: overview and perspectives for Brazil*. Rev Saude Publica, 2016. **50**: p. 17.
135. Kullenberg C, Kasperowski D. *What Is Citizen Science?--A Scientometric Meta-Analysis*. PLoS One, 2016. **11**(1): p. e0147152.
136. Smolinski MS, et al. *Participatory Disease Surveillance: Engaging Communities Directly in Reporting, Monitoring, and Responding to Health Threats*. JMIR Public Health Surveill, 2017. **3**(4): p. e62.

137. Chunara R., Smolinski MS, Brownstein JS. *Why we need crowdsourced data in infectious disease surveillance*. *Curr Infect Dis Rep*, 2013. **15**(4): p. 316-9.
138. Paolotti D, et al. *Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience*. *Clin Microbiol Infect*, 2014. **20**(1): p. 17-21.
139. Guerrisi C, et al. *Participatory Syndromic Surveillance of Influenza in Europe*. *J Infect Dis*, 2016. **214**(suppl_4): p. S386-s392.
140. Koppeschaar CE, et al. *Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe*. *JMIR Public Health Surveill*, 2017. **3**(3): p. e66.
141. Marquet RL, et al, *Internet-based monitoring of influenza-like illness (ILI) in the general population of the Netherlands during the 2003-2004 influenza season*. *BMC Public Health*, 2006. **6**: p. 242.
142. van Noort SP, et al. *Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe*. *Euro Surveill*, 2007. **12**(7): p. E5-6.
143. Perrotta D, et al. *Participatory Online Surveillance as a Supplementary Tool to Sentinel Doctors for Influenza-Like Illness Surveillance in Italy*. *PLoS One*, 2017. **12**(1): p. e0169801.
144. Adler AJ, et al. *Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey*. *BMC Infect Dis*, 2014. **14**: p. 232.
145. Pini A, et al. *High added value of a population-based participatory surveillance system for community acute gastrointestinal, respiratory and influenza-like illnesses in Sweden, 2013-2014 using the web*. *Epidemiol Infect*, 2017. **145**(6): p. 1193-1202.
146. Rehn M, et al. *Evaluation of an Internet-based monitoring system for influenza-like illness in Sweden*. *PLoS One*, 2014. **9**(5): p. e96740.
147. Debin M, et al. *Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France*. *PLoS One*, 2013. **8**(9): p. e73675.
148. Guerrisi C, et al. *The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France*. *Euro Surveill*, 2018. **23**(25).
149. Kjelso C, et al. *Influmeter - an online tool for self-reporting of influenza-like illness in Denmark*. *Infect Dis (Lond)*, 2016. **48**(4): p. 322-327.
150. Dalton C, et al. *Insights From Flutracking: Thirteen Tips to Growing a Web-Based Participatory Surveillance System*. *JMIR Public Health Surveill*, 2017. **3**(3): p. e48.
151. Chunara R, et al. *Estimating influenza attack rates in the United States using a participatory cohort*. *Sci Rep*, 2015. **5**: p. 9540.
152. Smolinski MS, et al. *Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons*. *Am J Public Health*, 2015. **105**(10): p. 2124-30.
153. Bayer C, et al, *Internet-based syndromic monitoring of acute respiratory illness in the general population of Germany, weeks 35/2011 to 34/2012*. *Euro Surveill*, 2014. **19**(4).
154. Leal Neto O, et al. *Saude na Copa: The World's First Application of Participatory Surveillance for a Mass Gathering at FIFA World Cup 2014, Brazil*. *JMIR Public Health Surveill*, 2017. **3**(2): p. e26.
155. Yano T, et al. *A Participatory System for Preventing Pandemics of Animal Origins: Pilot Study of the Participatory One Health Disease Detection (PODD) System*. *JMIR Public Health Surveill*, 2018. **4**(1): p. e25.
156. Lwin MO, et al. *Lessons From the Implementation of Mo-Buzz, a Mobile Pandemic Surveillance System for Dengue*. *JMIR Public Health Surveill*, 2017. **3**(4): p. e65.
157. Karimuribo ED, et al. *A Smartphone App (AfyaData) for Innovative One Health Disease Surveillance from Community to National Levels in Africa: Intervention in Disease Surveillance*. *JMIR Public Health Surveill*, 2017. **3**(4): p. e94.

158. Schmidt C, et al. *Kidenga: Public engagement for detection and prevention of Aedes-borne viral diseases*. Online J Public Health Inform, 2017. **9**(1): p. e111.
159. Genevieve LD, et al. *Participatory Disease Surveillance Systems: Ethical Framework*. J Med Internet Res, 2019. **21**(5): p. e12273.
160. ECDC. *EU case definitions*. Available from: <https://ecdc.europa.eu/en/surveillance-and-disease-data/eu-case-definitions>.
161. Cantarelli P, et al. *The representativeness of a European multi-center network for influenza-like-illness participatory surveillance*. BMC Public Health, 2014. **14**: p. 984.
162. Fox S, Rainie L. *E-patients and the online health care revolution*. Physician Exec, 2002. **28**(6): p. 14-7.
163. Vandendijck Y, Faes C, Hens N. *Eight years of the Great Influenza Survey to monitor influenza-like illness in Flanders*. PLoS One, 2013. **8**(5): p. e64156.
164. Friesema IH, et al. *Internet-based monitoring of influenza-like illness in the general population: experience of five influenza seasons in The Netherlands*. Vaccine, 2009. **27**(45): p. 6353-7.
165. Kalimeri K, et al. *Unsupervised extraction of epidemic syndromes from participatory influenza surveillance self-reported symptoms*. PLoS Comput Biol, 2019. **15**(4): p. e1006173.
166. van Noort SP, et al. *Ten-year performance of Influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour*. Epidemics, 2015. **13**: p. 28-36.
167. Eames KT, et al. *Rapid assessment of influenza vaccine effectiveness: analysis of an internet-based cohort*. Epidemiol Infect, 2012. **140**(7): p. 1309-15.
168. Debin M, et al. *Effectiveness of 2012-2013 influenza vaccine against influenza-like illness in general population: estimation in a French web-based cohort*. Hum Vaccin Immunother, 2014. **10**(3): p. 536-43.
169. Ariza M, et al. *Healthcare-seeking behaviour in case of influenza-like illness in the French general population and factors associated with a GP consultation: an observational prospective study*. BJGP Open, 2018. **1**(4): p. bjgpopen17X101253.
170. Van Cauteren D, et al. *Burden of influenza, healthcare seeking behaviour and hygiene measures during the A(H1N1)2009 pandemic in France: a population based study*. BMC Public Health, 2012. **12**: p. 947.
171. Guerrisi C, et al. *Factors associated with influenza-like-illness: a crowdsourced cohort study from 2012/13 to 2017/18*. BMC Public Health, 2019. **19**(1): p. 879.
172. Bajardi P, et al. *Determinants of follow-up participation in the Internet-based European influenza surveillance platform Influenzanet*. J Med Internet Res, 2014. **16**(3): p. e78.
173. Baltrusaitis K, et al. *Determinants of Participants' Follow-Up and Characterization of Representativeness in Flu Near You, A Participatory Disease Surveillance System*. JMIR Public Health Surveill, 2017. **3**(2): p. e18.
174. Bajardi P, et al. *Association between recruitment methods and attrition in Internet-based studies*. PLoS One, 2014. **9**(12): p. e114925.
175. Larson HJ, et al. *Addressing the vaccine confidence gap*. Lancet, 2011. **378**(9790): p. 526-35.
176. Schwarzinger M, et al. *Low acceptability of A/H1N1 pandemic vaccination in French adult population: did public health policy fuel public dissonance?* PLoS One, 2010. **5**(4): p. e10199.
177. Larson HJ, et al. *The State of Vaccine Confidence 2016: Global Insights Through a 67-Country Survey*. EBioMedicine, 2016. **12**: p. 295-301.
178. Boiron K, et al. *Opinion about seasonal influenza vaccination among the general population 3 years after the A(H1N1)pdm2009 influenza pandemic*. Vaccine, 2015. **33**(48): p. 6849-54.
179. Loubet P, et al. *First nationwide web-based surveillance system for influenza-like illness in pregnant women: participation and representativeness of the French G-GrippeNet cohort*. BMC Public Health, 2016. **16**: p. 253.

180. Loubet P, et al. *Influenza during pregnancy: Incidence, vaccination coverage and attitudes toward vaccination in the French web-based cohort G-GrippeNet*. *Vaccine*, 2016. **34**(20): p. 2390-6.
181. Quade P, Nsoesie EO. *A Platform for Crowdsourced Foodborne Illness Surveillance: Description of Users and Reports*. *JMIR Public Health Surveill*, 2017. **3**(3): p. e42.
182. Wenham C, et al. *Self-Swabbing for Virological Confirmation of Influenza-Like Illness Among an Internet-Based Cohort in the UK During the 2014-2015 Flu Season: Pilot Study*. *J Med Internet Res*, 2018. **20**(3): p. e71.
183. Haussig JM, et al. *Feasibility study for the use of self-collected nasal swabs to identify pathogens among participants of a population-based surveillance system for acute respiratory infections (GrippeWeb-Plus)-Germany, 2016*. *Influenza Other Respir Viruses*, 2019. **13**(4): p. 319-330.
184. Myers J, et al. *Ethics in public health research: privacy and public health at risk: public health confidentiality in the digital age*. *Am J Public Health*, 2008. **98**(5): p. 793-801.
185. Bousquet V, Caserio-Schönemann C. *La surveillance des urgences par le réseau OSCOUR (Organisation de la surveillance coordonnée des urgences)*. 2012: p. 12 p.
186. World Health Organization. *WHO Regional Office for Europe guidance for sentinel influenza surveillance in humans*. 2011: Geneva.
187. Vega T, et al. *Influenza surveillance in Europe: establishing epidemic thresholds by the moving epidemic method*. *Influenza Other Respir Viruses*, 2013. **7**(4): p. 546-58.
188. Vega T, et al. *Influenza surveillance in Europe: comparing intensity levels calculated using the moving epidemic method*. *Influenza Other Respir Viruses*, 2015. **9**(5): p. 234-46.
189. Harcourt SE, et al. *Developing influenza and respiratory syncytial virus activity thresholds for syndromic surveillance in England*. *Epidemiol Infect*, 2019. **147**: p. e163.
190. Campbell H, Bont L, Nair H. *Respiratory syncytial virus (RSV) disease - new data needed to guide future policy*. *J Glob Health*, 2015. **5**(2): p. 020101.
191. Thompson WW, et al. *Influenza-associated hospitalizations in the United States*. *Jama*, 2004. **292**(11): p. 1333-40.
192. Martinez A, et al. *Risk factors associated with severe outcomes in adult hospitalized patients according to influenza type and subtype*. *PLoS One*, 2019. **14**(1): p. e0210353.
193. Viboud C, et al. *Risk factors of influenza transmission in households*. *Br J Gen Pract*, 2004. **54**(506): p. 684-9.
194. Ferguson NM, et al. *Strategies for mitigating an influenza pandemic*. *Nature*, 2006. **442**(7101): p. 448-52.
195. Klick B, Nishiura H, Cowling BJ. *Optimal design of intervention studies to prevent influenza in healthy cohorts*. *PLoS One*, 2012. **7**(4): p. e35166.
196. Cauchemez S, et al. *A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data*. *Stat Med*, 2004. **23**(22): p. 3469-87.
197. Petrie JG, et al. *Influenza transmission in a cohort of households with children: 2010-2011*. *PLoS One*, 2013. **8**(9): p. e75339.
198. Cowling BJ, et al. *Facemasks and hand hygiene to prevent influenza transmission in households: a cluster randomized trial*. *Ann Intern Med*, 2009. **151**(7): p. 437-46.
199. Delabre RM, et al. *Risk factors of pandemic influenza A/H1N1 in a prospective household cohort in the general population: results from the CoPanFlu-France cohort*. *Influenza Other Respir Viruses*, 2015. **9**(1): p. 43-50.
200. Pebody RG, et al. *Assessing the severity of influenza: a role for longitudinal telephone surveys?* *Epidemiol Infect*, 2018. **146**(16): p. 2042-2048.
201. von der Beck D, et al. *Characteristics and outcomes of a cohort hospitalized for pandemic and seasonal influenza in Germany based on nationwide inpatient data*. *PLoS One*, 2017. **12**(7): p. e0180920.

202. Chowell G, et al. *Characterizing the epidemiology of the 2009 influenza A/H1N1 pandemic in Mexico*. PLoS Med, 2011. **8**(5): p. e1000436.
203. Lemaitre M, Carrat F. *Comparative age distribution of influenza morbidity and mortality during seasonal influenza epidemics and the 2009 H1N1 pandemic*. BMC Infect Dis, 2010. **10**: p. 162.
204. Klein SL, Hodgson A, Robinson DP. *Mechanisms of sex disparities in influenza pathogenesis*. J Leukoc Biol, 2012. **92**(1): p. 67-73.
205. Nicoll A, et al. *Experience and lessons from surveillance and studies of the 2009 pandemic in Europe*. Public Health, 2010. **124**(1): p. 14-23.
206. Ung A, et al. *Disentangling a complex nationwide Salmonella Dublin outbreak associated with raw-milk cheese consumption, France, 2015 to 2016*. Euro Surveill, 2019. **24**(3).
207. Mathieu P, et al. *Population perception of mandatory childhood vaccination programme before its implementation, France, 2017*. Euro Surveill, 2019. **24**(25).
208. Nicolay N, et al. *Mandatory immunization: the point of view of the French general population and practitioners*. Vaccine, 2008. **26**(43): p. 5484-93.
209. Reich NG, et al. *A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States*. Proc Natl Acad Sci U S A, 2019. **116**(8): p. 3146-3154.
210. Santillana M, et al. *Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance*. PLoS Comput Biol, 2015. **11**(10): p. e1004513.
211. McGowan CJ, et al. *Collaborative efforts to forecast seasonal influenza in the United States, 2015-2016*. Sci Rep, 2019. **9**(1): p. 683.
212. Vergu E, et al. *Medication sales and syndromic surveillance, France*. Emerg Infect Dis, 2006. **12**(3): p. 416-21.
213. Ertem Z, Raymond D, Meyers LA. *Optimal multi-source forecasting of seasonal influenza*. PLoS Comput Biol, 2018. **14**(9): p. e1006236.

ANNEXE 1. Questionnaire d'inclusion

Les questions en rouge sont celles ajoutées spécifiquement à la cohorte GrippeNet.fr par rapport aux autres pays européens.

Les questions précédées d'un astérisque sont des questions nécessitant une réponse obligatoirement.

Contexte Q0 (cette question est posée lors de l'ajout d'un nouveau participant. Pour chaque nouveau participant cette question est posée) ***Pour qui remplissez-vous ce questionnaire ?**

- Pour moi-même
- Pour une personne âgée de moins de 18 ans dont je suis le parent ou représentant légal
- Pour un membre majeur de mon foyer
- Pour quelqu'un d'autre

Contexte Q23 (si réponse à la Q0a est "Pour une personne âgée de moins de 18 ans dont je suis le parent ou représentant légal") ***Je certifie avoir expliqué à cette personne le contenu de cette étude et avoir obtenu son libre consentement pour y participer. Elle m'a donné pouvoir pour remplir les questionnaires à sa place. Je certifie être l'un de ses représentants légaux.**

- Oui
- Non

Si la case « non » est cochée, il n'est pas possible de poursuivre le remplissage du questionnaire pour cette personne. Il n'est également pas possible de remplir les autres questionnaires. Si la case oui est cochée, le texte suivant apparaît : « Répondez s'il-vous-plaît à toutes les questions comme si vous étiez la personne pour laquelle vous remplissez ce questionnaire. ».

Contexte Q22 (si réponse à la Q0a est "Pour un membre de mon foyer" ou "Pour quelqu'un d'autre") **Merci de demander à cette personne de répondre à la question : « J'ai pris connaissance du contenu de cette étude et souhaite librement y participer. Je donne pouvoir à la personne détentrice de ce compte pour remplir mes questionnaires à ma place. »**

- Oui
- Non

Si la case « non » est cochée, il n'est pas possible de poursuivre le remplissage du questionnaire. Il n'est également pas possible de remplir les autres questionnaires. Si la case oui est cochée, le texte suivant apparaît : « Répondez s'il-vous-plaît à toutes les questions comme si vous étiez la personne pour laquelle vous remplissez ce questionnaire. ».

NOTE 2 : Ce premier questionnaire permet de mieux connaître le profil des participants de GrippeNet.fr, et vous prendra entre 5 et 10 minutes. Son remplissage n'a lieu qu'une fois par saison et sera pré-complété si vous avez déjà participé les années passées. Vous pourrez revenir sur le questionnaire pour nous indiquer un changement de situation (situation professionnelle, statut vaccinal, etc.)

Contexte Q1 ***Quel est votre sexe ?**

- Homme
- Femme

Contexte Q2 ***Quelle est votre date de naissance (mois et année)?**

Si la réponse à Q0 est « pour moi-même » et que la date de naissance correspond à un individu âgé de moins de 18 ans, un message s'affiche pour demander à ce que cela soit les parents du mineur qui répondent aux premières questions).

Contexte Q19 **Quelle est votre taille (en centimètres)?**

Contexte Q20 **Quelle est votre poids actuel (en kilogrammes) ?**

Contexte Q12 (posée uniquement aux femmes entre 15 et 50 ans) **Êtes-vous enceinte ?**

- Oui [renvoi à QG2]
- Non [renvoi à Q13]

- Je ne sais pas / préfère ne pas répondre [renvoi à Q13]

Contexte Q3 *Quel est la commune de votre domicile ?

Contexte Q4 *Quelle est votre activité principale actuellement?

- Salarié(e), à temps plein
- Salarié(e), à temps partiel
- Travailleur (se) indépendant(e) (entrepreneur, commerçant, agriculteur, profession libérale...)
- Etudiant(e), écolier ou enfant gardé en dehors du domicile (ex : à la crèche)
- Homme ou femme au foyer
- Demandeur(se) d'emploi
- En congé maladie longue durée ou en congé parental
- Retraité(e)
- Aucune de ces réponses

Contexte Q4b (si « emploi rémunéré », « travailleur indépendant » ou « écolier / étudiant » à Q4) Quel est la commune de votre école / université / lieu de travail (là où vous passez la plus grande partie de votre temps à travailler/étudier)?

- XXXX (il faut entrer le code postal de la commune et sélectionner la commune correspondante)
- Je ne sais pas / ne me souviens pas
- Non applicable (par exemple, je n'ai pas de lieu de travail fixe)

Contexte Q4c (si « Salarié temps plein » ou « Salarié temps partiel » à Q4) Champs libre si « Autre » est coché. Laquelle de ces propositions correspond le plus à votre activité principale :

- Cadre ou profession intermédiaire n'exerçant pas un travail de bureau (médecin, professeur, infirmier...)
- Cadre ou profession intermédiaire exerçant un travail de bureau (administrateur, comptable...)
- Profession dans le domaine du commerce, de la restauration, de l'hôtellerie, du loisir (vendeur, serveur, professeur de sport...)
- Ouvrier qualifié, artisan (mécanicien, électricien, technicien...)
- Autre métier manuel (agent d'entretien, agent de sécurité, chauffeur...)
- Autre

Contexte Q4d (question posée uniquement aux personnes âgées de 16 ans et plus) Quel est votre plus haut niveau de qualification ? Si vous êtes toujours scolarisé ou étudiant, cochez s'il-vous-plaît la case correspondant au diplôme le plus élevé que vous avez obtenu jusqu'à maintenant.

- o Je n'ai aucun diplôme
- o Diplôme national du brevet, CAP, BEP, CFPA...
- o Baccalauréat, brevet professionnel, brevet de technicien...
- o Licence, DUT, BTS...
- o Master, diplôme d'ingénieur ou diplôme de troisième cycle (thèse d'université ou d'exercice...)
- o Je suis toujours scolarisé ou étudiant

Contexte Q5 *Sans compter les personnes que vous croisez dans les lieux publics (transports en commun, magasins...), avez-vous des contacts avec les personnes suivantes au cours d'une journée typique ? (Sélectionnez toutes les options qui s'appliquent, s'il y a lieu)

- o Plus de 10 enfants ou adolescents dans une même journée
- o Entre 1 et 9 enfants ou adolescents dans la même journée (en ne comptant pas vos propres enfants)
- o Plus de 10 personnes âgées de plus de 65 ans dans une même journée
- o Des personnes malades
- o Des groupes de personnes (partage d'un même lieu avec plus de 10 individus en même temps, ex : dans un bureau)
- o Aucune de ces réponses

Contexte Q6 En dehors de vous, combien de personnes de chacune des classes d'âge suivantes vivent dans votre foyer ? Ne vous comptez pas dans l'une de ces catégories, indiquez seulement le nombre d'autres personnes de votre foyer pour chaque catégorie cochée.

Je vis seul(e)

- 0 - 4 ans
- 5 - 18 ans
- 19 - 44 ans
- 45 - 64 ans

65 ans et +

Contexte Q6b (*uniquement posée s'il y a au moins un enfant de 0-4 ans dans le foyer, y compris des participants*)
Combien d'enfants de votre foyer sont gardés en dehors du domicile (école, crèche, garderie, assistante maternelle...)?

Aucun

1

2

3

4

5

>5

Contexte Q21 (*posée si réponse à la Q6b*) **Dans quelle(s) structure(s) ce ou ces enfant(s) sont-ils gardés ?**

Ecole

Assistante maternelle

Garde partagée

Crèche (collective, parentale, micro-crèche)

Crèche familiale

Jardin d'enfant

Halte-garderie

Contexte Q16 **Avez-vous des animaux à la maison ?** (Sélectionnez toutes les options qui s'appliquent)

Non

Oui, un ou plusieurs chiens

Oui, un ou plusieurs chats

Oui, un ou plusieurs oiseaux

Oui, un ou plusieurs autres animaux

Contexte Q7 * **Quel est votre principal moyen de déplacement ?**

• Marche

• Vélo

• Moto / scooter

• Voiture

• Transports en commun (bus, train, métro, etc.)

• Autre

Contexte Q7b * **Lors d'une journée ordinaire, combien de temps passez-vous au total dans les transports en commun (bus, train, métro, etc.)?**

• Pas de temps du tout

• 0 - 30 minutes

• 30 minutes - 1 heure et demie

• 1 heure et demie - 4 heures

• Plus de 4 heures

Contexte Q8 **A quelle fréquence avez-vous des rhumes ou la grippe ?**

• Presque jamais

• Parfois, pas tous les ans

• Une ou deux fois par an

• Entre 3 et 5 fois par an

• Entre 6 et 10 fois par an

• Plus de 10 fois par an

• Je ne sais pas

Contexte Q18 * **Avez-vous eu un bon de vaccination contre la grippe cette année (depuis septembre 2018)?**

• Oui

• Non

• Je ne sais pas / ne me souviens pas

Contexte Q10 * **Avez-vous été vacciné(e) contre la grippe cette année (depuis le mois de septembre 2018) ?**

- Oui [renvoi à Q9b et Q9c]
- Non [renvoi à Q9d]
- Je ne sais pas / ne me souviens pas

Contexte Q10b (si « Oui » à la Q10) **Quand avez-vous été vacciné(e) contre la grippe cette année (depuis septembre 2018)?**

- Choisissez la date : XX / XX / XXXX
- Je ne sais pas / ne me souviens pas

Contexte Q10e (si « Oui » à la Q10) **Par qui avez-vous été vacciné ?**

- Par un(e) médecin généraliste
- Par un(e) infirmier(e)
- Par une sage-femme
- Par un(e) pharmacien(ne)
- Par un(e) médecin ou infirmier(e) du travail
- Par un(e) professionnel(le) d'une autre spécialité (préciser)(champs libre)

Contexte Q10c (si « Oui » à la Q10) **Quelles étaient vos raisons pour vous faire vacciner contre la grippe cette année ?** (Sélectionnez toutes les options qui s'appliquent)

- J'ai eu un bon de vaccination
- Je fais partie d'un groupe pour lequel la vaccination est recommandée (femmes enceintes, plus de 65 ans, problèmes de santé sous-jacents, contacts avec des enfants de moins de 6 mois à risque, résidents dans un établissement de soins de longue durée ou médico-social, obèses, professionnels de santé, etc.)
- Je suis enceinte et je souhaite protéger mon bébé contre la grippe
- La vaccination diminue mes risques de contracter la grippe
- La vaccination diminue le risque de transmettre la grippe à d'autres personnes
- Mon médecin me l'a recommandé
- Cela m'a été recommandé dans mon lieu de travail / école
- Le vaccin était aisément disponible et son administration était pratique
- Le vaccin était gratuit
- Je ne veux pas manquer le travail / l'école
- Je me fais toujours vacciner
- Autre(s) raison(s)

Contexte Q10d (si « Non » à la Q10) **Pourquoi n'avez-vous pas été vacciné(e) contre la grippe saisonnière cette année ?** (Sélectionnez toutes les options qui s'appliquent)

- Je prévois de me faire vacciner, mais je ne l'ai pas encore fait
- La vaccination ne m'a pas été proposée
- On m'a déconseillé de me faire vacciner car je suis enceinte
- Je suis enceinte et je crains que le vaccin présente un danger pour mon bébé
- Je n'appartiens pas à un groupe pour lequel la vaccination est recommandée (femmes enceintes, plus de 65 ans, problèmes de santé sous-jacents, contacts avec des enfants de moins de 6 mois à risque, résidents dans un établissement de soins de longue durée ou médico-social, obèses, professionnels de santé, etc.)
- Il est préférable de se construire sa propre immunité naturelle contre la grippe
- Je doute que le vaccin contre la grippe soit efficace
- La grippe est une maladie bénigne
- Je ne pense pas que je sois susceptible de contracter la grippe
- Je pense que le vaccin contre la grippe peut causer la grippe
- Je crains que le vaccin ne soit pas sûr ou cause une maladie ou d'autres événements indésirables
- Je n'aime pas recevoir de vaccins
- Le vaccin n'est pas facilement disponible pour moi
- Le vaccin n'est pas gratuit
- La vaccination m'a été recommandée par mon médecin, mais je ne me suis pas fait vacciner
- Je ne sais pas
- Autre(s) raison(s)

Contexte Q9 *L'an dernier (en 2017-2018) avez-vous été vacciné(e) contre la grippe ?

- Oui
- Non
- Je ne sais pas / ne me souviens pas

Contexte Q11 *Prenez-vous régulièrement des médicaments pour l'un des problèmes médicaux suivants ?

(Sélectionnez toutes les options qui s'appliquent)

- Non, je n'ai aucune des maladies qui figurent ci-dessous
- Asthme
- Diabète
- Troubles pulmonaires (BPCO, emphysème...) ou autres désordres affectant votre respiration
- Troubles cardiaques
- Troubles rénaux
- Immunodépression (ablation de la rate, maladie du foie, greffe d'organe, immunodéficience acquise dont VIH, traitement de cancer en cours, prise de corticoïde ou d'un traitement immunosuppresseur...)

Contexte Q13 * Fumez-vous ?

- Non, je n'ai jamais fumé
- Non, j'ai arrêté de fumer depuis plus d'un an
- Non, j'ai arrêté de fumer depuis moins d'un an
- Oui, occasionnellement
- Oui, tous les jours, moins de 10 fois par jour
- Oui, tous les jours, 10 fois ou plus par jour
- Je ne sais pas / préfère ne pas répondre

Contexte Q14 *Avez-vous l'une des allergies respiratoires suivantes ? (Sélectionnez toutes les options qui s'appliquent)

- Rhume des foins
- Allergie aux acariens
- Allergie aux animaux domestiques
- Autres allergies provoquant des symptômes respiratoires (par exemple des éternuements, des yeux larmoyants) (*champs libre*)
- Je n'ai pas d'allergie causant des symptômes respiratoires

Contexte Q24 Prenez-vous ou prévoyez-vous de prendre cet hiver des médicaments homéopathiques pour vous protéger de la grippe ?

- Oui
- Non
- Je ne sais pas
- Je ne souhaite pas répondre

Contexte Q17 Comment avez-vous eu connaissance de l'étude GrippeNet.fr ?

- Par la télévision ou la radio
- Par un journal ou magazine
- Par un site Internet institutionnel (Santé Publique France, Ameli...)
- Par un site Internet d'informations généralistes (type journaux en ligne)
- Par un site Internet d'information en santé
- Par un email
- Par un réseau social
- Par une affiche du projet
- Par un membre de ma famille ou un ami
- A l'école ou au travail
- Par un professionnel de la santé
- Par un de mes proches qui travaille sur le projet
- Autre (préciser)

ANNEXE 2. Questionnaire hebdomadaire de symptômes

Les questions en rouge sont celles ajoutées spécifiquement à la cohorte GrippeNet.fr par rapport aux autres pays européens.

Les questions précédées d'un astérisque sont des questions nécessitant une réponse obligatoirement.

Q1 hebdomadaire *Depuis votre dernière connexion (ou dans les sept derniers jours, si cela est votre première visite de la saison), avez-vous eu certains des symptômes suivants ? (Sélectionnez toutes les options qui s'appliquent)

- Pas de symptôme
- Fièvre
- Frissons
- Nez qui coule ou bouché
- Éternuements
- Maux de gorge
- Toux
- Essoufflement
- Maux de tête
- Douleurs musculaires / articulaires
- Douleur thoracique
- Fatigue ou épuisement (malaise)
- Perte d'appétit
- Expectorations (crachats) colorées
- Yeux larmoyants ou injectés de sang
- Nausées
- Vomissements
- Diarrhée
- Maux d'estomac
- Autre

Q2 hebdomadaire (Si le participant était encore malade lors de sa dernière visite et a signalé des symptômes aujourd'hui) Lors de votre dernière connexion le « date de la dernière connexion », vous avez indiqué que vous étiez malade depuis le « date de début des symptômes ». Les symptômes que vous rapportez aujourd'hui sont-ils la suite de la maladie de la dernière fois ?

- Oui
- Non
- Je ne sais pas / ne me souviens pas

Q15 hebdomadaire (si symptômes)(si réponse à Q2 autre que « Oui ») Au moment où les symptômes ont débuté, étiez-vous :

- Dans votre département de résidence habituel
- En France métropolitaine hors de votre département de résidence habituel
- A l'étranger (ou en DOM TOM)

Q5a hebdomadaire (si diarrhée) Combien de selles par jour avez-vous ?

- Moins de 3
- Plus de 3

Q3 hebdomadaire (si symptômes) Quand les premiers symptômes sont-ils apparus ?

- Choisissez la date XX / XX / XXXX
- Je ne sais pas / ne me souviens pas

Q4 hebdomadaire (si symptômes) Quand se sont-ils terminés ?

- Choisissez la date XX / XX / XXXX

- Je ne sais pas / ne me souviens pas
- Je suis encore malade

Q5 hebdomadaire (*si symptômes*) **Les symptômes se sont-ils développés subitement en quelques heures ?**

- Oui
- Non
- Je ne sais pas / ne me souviens pas

Q6 hebdomadaire (*si fièvre*) **Quand la fièvre a-t-elle commencé ?**

- Choisissez la date XX / XX / XXXX
- Je ne sais pas / ne me souviens pas

Q6b hebdomadaire (*si fièvre*) **La fièvre est-elle apparue subitement en quelques heures ?**

- Oui
- Non
- Je ne sais pas

Q6c hebdomadaire (*si symptômes*) **Avez-vous pris votre température ?**

- Oui [*renvoi à Q6c hebdomadaire*]
- Non
- Je ne sais pas

Q6d hebdomadaire (*si symptômes*) et (*si prise de température*) **Quelle a été votre plus haute température mesurée ?**

- Inférieure à 37°C
- 37° - 37,4°C
- 37,5° - 37,9°C
- 38° - 38,9°C
- 39° - 39,9°C
- 40°C ou plus
- Je ne sais pas / ne me souviens pas

Q8 hebdomadaire (*si symptômes*) **En raison de vos symptômes, avez-vous contacté par téléphone un médecin généraliste ?**

- Oui – je n'ai pas parlé directement au médecin
- Oui – j'ai parlé directement au médecin
- Non
- Autre

Q8b hebdomadaire (*si symptômes et si réponses 1 ou 2 à la Q8*) **Combien de temps après l'apparition des symptômes avez-vous contacté par téléphone ce médecin pour la première fois ?**

Médecin généraliste (je n'ai pas parlé directement au médecin)

Médecin généraliste (j'ai parlé directement au médecin)

Autre

Pour chaque possibilité, menu déroulant avec option à choisir :

- Le jour même
- 1 jour après
- 2 jours
- 3 jours
- 4 jours
- 5 - 7 jours
- Plus de 7 jours après
- Je ne sais pas / ne me souviens pas

Q7 hebdomadaire (*si symptômes*) **En raison de vos symptômes, avez-vous consulté (vu en personne) depuis votre dernière connexion un des services médicaux suivants ? (Sélectionnez toutes les options qui s'appliquent)**

- Non, aucun
- Non, mais j'ai un rendez-vous
- Médecin généraliste
- Autre médecin de ville (pédiatre, ORL, cardiologue...)

- o Gynécologue / Obstétricien
- o Sage-femme
- o Pharmacien
- o Infirmière scolaire (pour les -18 ans seulement)
- o Urgences hospitalières / service de garde / SOS Médecins
- o Autres services médicaux (infirmière, médecine du travail, PMI...)

Q7b hebdomadaire (si symptômes) - cela est demandé pour chaque item coché lors de la question précédente (sauf pour «Aucun» ou «Non, mais j'ai un rendez-vous») **Combien de temps après l'apparition des symptômes avez-vous consulté pour la première fois ?**

Médecin généraliste

Autre médecin de ville (gynécologue, pédiatre, ORL, cardiologue...)

Urgences hospitalières / service de garde / SOS Médecins

Autres services médicaux (infirmière, médecine du travail, sage-femme, PMI, pharmacien...)

Pour chaque possibilité, menu déroulant avec option à choisir :

- Le jour même
- 1 jour après
- 2 jours
- 3 jours
- 4 jours
- 5 - 7 jours
- Plus de 7 jours après
- Je ne sais pas / ne me souviens pas

Q9 hebdomadaire (si symptômes) **Avez-vous pris des médicaments pour ces symptômes ?** (Sélectionnez toutes les options qui s'appliquent)

- o Non
- o Médicaments contre la douleur ou la fièvre (par exemple paracétamol, ibuprofène, aspirine, etc.)
- o Médicaments contre la toux (par exemple expectorants)
- o Antiviraux (Tamiflu, Relenza)
- o Antibiotiques
- o Homéopathie
- o Médecines douces (huiles essentielles, phytothérapie, etc.)
- o Autres (précisez)
- o Je ne sais pas / ne me souviens pas

Q9c hebdomadaire (si "Antiviraux" a été coché) **Concernant les antiviraux que vous avez pris :**

- o Ils viennent de vous être prescrits par un médecin pour les symptômes que vous rapportez aujourd'hui
- o Vous les aviez déjà chez vous

Q9b hebdomadaire (si des antiviraux ont été pris) **Combien de temps après le début de vos symptômes avez-vous commencé à prendre des médicaments antiviraux ?**

- Le même jour (dans les 24 heures)
- 1 jour après
- 2 jours
- 3 jours
- 4 jours
- 5 - 7 jours
- Plus de 7 jours après
- Je ne sais pas / ne me souviens pas

Q9d hebdomadaire (si "Antibiotiques" a été coché) **Concernant les antibiotiques que vous avez pris :**

- o Ils viennent de vous être prescrits par un médecin pour les symptômes que vous rapportez aujourd'hui
- o Vous les aviez déjà chez vous

Q14 hebdomadaire (si symptômes) **Avez-vous été hospitalisé(e) à cause des symptômes que vous rapportez aujourd'hui ?**

- Oui
- Non

Q8c hebdomadaire (si symptômes) **En raison de vos symptômes, avez-vous consulté un site internet ?**

- Oui
- Non

Q10 hebdomadaire (si symptômes) **Avez-vous changé vos habitudes de vie à cause de votre maladie (ex : annuler une activité sportive ou culturelle, sauter un repas...)?**

- Oui, mais je n'ai pas eu d'arrêt de travail / d'école
- Oui, et j'ai eu un arrêt de travail / d'école
- Non

Q10b hebdomadaire (si symptômes) et (si arrêt maladie Q10) **Etes-vous toujours en arrêt maladie / d'école ?**

- Oui
- Non
- Autres (par exemple, je n'avais pas à être au travail aujourd'hui)

Q10c hebdomadaire (si symptômes et si arrêt maladie Q10) **Combien de temps avez-vous été en arrêt maladie / d'école (si vous êtes toujours en arrêt : depuis combien de temps l'êtes-vous)?**

- 1 jour
- 2 jours
- 3 jours
- 4 jours
- 5 jours
- 6 - 10 jours
- 11 - 15 jours
- Plus de 15 jours

Q11 hebdomadaire (si symptômes) **Quelle est selon vous l'origine de vos symptômes ?**

- Grippe ou syndrome grippal
- Mon médecin m'a dit qu'il s'agissait de la grippe
- Rhume
- Allergie / rhume des foins
- Gastro-entérite / grippe intestinale
- Asthme
- Autre (maladie diagnostiquée par le médecin)
- Autre
- Je ne sais pas

Q16 hebdomadaire **Globalement, comment vous sentez-vous ?**

Echelle de 0 à 100 (0 : Très mal ; 100 : Tout va bien)