



MASTER 2<sup>ème</sup> ANNÉE  
Santé publique Paris Sud-Saclay et Santé UPEC

Dominante  
**ÉPIDEMIOLOGIE ET SURVEILLANCE DES MALADIES INFECTIEUSES  
HUMAINES ET ANIMALES**

---

**RAPPORT DE STAGE**

**ANALYSE DE L'EXHAUSTIVITÉ DE LA SURVEILLANCE DES CAS GRAVES DE GRIPPE ADMIS EN  
RÉANIMATION, SAISON 2021-2022**

**Présenté par**

Intila YALIKUN

Réalisé sous la direction de : Sibylle BERNARD-STOECKLIN

Organisme et pays : Santé publique France (Saint-Maurice), France

Période du stage : 2 janvier 2023 au 23 juin 2023

Date de soutenance : 26 juin 2023

Année universitaire 2022-2023

## Remerciements

J'aimerais tout d'abord adresser ma gratitude à Sibylle BERNARD-STOECKLIN pour sa bienveillance, son écoute et sa disponibilité tout au long de ce stage. J'ai beaucoup apprécié travailler avec toi. Merci de m'avoir permis de réaliser ce stage.

J'aimerais également adresser mes sincères remerciements à Mathilde PIVETTE, Cécile SOMMEN, Thibault BOUDON, Cécile DURAND, Isabelle PARENT DU CHÂTELET, Daniel LEVY-BRUHL, Laurence MANDEREAU-BRUNO et Perrine DE CROUY-CHANEL pour leur accompagnement sur les aspects méthodologiques, que ce soit sur les étapes et modalités d'analyse, l'écriture du code, l'interprétation des résultats ou la rédaction du mémoire. Je remercie aussi Vincent AUVIGNE pour m'avoir aidé à rendre le code d'analyse disponible sur le Gitlab de Santé publique France.

Je remercie les cellules régionales de Santé publique France que nous avons sollicitées pour des questions portant sur les analyses régionales pour leur disponibilité.

Je tiens également à remercier l'unité REV pour leur accueil chaleureux dès le premier jour ainsi que tous mes voisins de bureau avec qui j'ai grandement apprécié discuter.

J'aimerais aussi adresser mes remerciements à l'équipe pédagogique du Master ESMIHA/SEMHA pour la qualité de leur accompagnement tout au long de l'année ainsi que leur disponibilité.

Enfin, j'aimerais remercier ma famille et mes amis pour leur soutien au quotidien.

## Liste des figures et des tableaux

### Liste des figures

<b>Figure 1</b> – Nombre de cas à chaque étape du nettoyage des bases de données SGS et PMSI préalable à l’analyse.....	<b>26</b>
<b>Figure 2</b> – Description et ordre d’exécution des requêtes de croisement.....	<b>28</b>
<b>Figure 3</b> – Incidence des cas graves de grippe admis en réanimation par semaine d’admission en réanimation, France entière, Semaine 40 de 2021 à Semaine 30 de 2022 .....	<b>33</b>
<b>Figure 4</b> – Répartition des cas par région d’admission classé par ordre décroissant, par base de données, France entière, Semaine 40 de 2021 à Semaine 30 de 2022.....	<b>34</b>
<b>Figure 5</b> – Diagramme de Venn avec le nombre de cas communs aux deux bases et le nombre de cas spécifique à chaque base.....	<b>36</b>
<b>Figure 6</b> – Répartition de l’incidence des cas graves de grippe admis en réanimation par semaine d’admission entre la métropole et les DROM, Semaine 40 de 2021 à Semaine 30 de 2022.....	<b>38</b>
<b>Figure 7</b> – Incidence des cas graves de grippe admis en réanimation par semaine d’admission, DROM, Semaine 40 de 2021 à Semaine 30 de 2022.....	<b>38</b>
<b>Figure 8</b> – Diagramme de Venn des effectifs selon le résultat du croisement de la base PMSI et SGS pour les cas graves de grippe admis en réanimation avec estimation du nombre de cas total selon l’estimateur de Chapman et de Chao, France entière, Semaine 40 de 2021 à Semaine 30 de 2022....	<b>39</b>
<b>Figure 9</b> – Estimation de l’exhaustivité par source au niveau national pour les analyses globales et restreintes aux services sentinelles, France entière, Semaine 40 de 2021 à Semaine 30 de 2022.....	<b>40</b>
<b>Figure 10</b> – Estimation de l’exhaustivité par source et par classe d’âge pour les analyses globales et restreintes aux services sentinelles, France entière, Semaine 40 de 2021 à Semaine 30 de 2022.....	<b>41</b>
<b>Figure 11</b> – Estimation de l’exhaustivité par source et par région pour les analyses globales et restreintes aux services sentinelles, par ordre décroissant de la valeur moyenne de l’exhaustivité SGS calculée avec l’estimateur de Chapman, France entière, Semaine 40 de 2021 à Semaine 30 de 2022	<b>42</b>
<b>Figure 12</b> – Carte du taux d’incidence pour 100 000 habitants des cas graves de grippe admis en réanimation estimé par région, à partir des valeurs moyennes de l’estimateur de Chapman (A) ou de Chao (B) issues de l’analyse globale.....	<b>45</b>

### Liste des tableaux

<b>Tableau 1</b> – Tableau de contingence du croisement de deux sources de données.....	<b>19</b>
<b>Tableau 2</b> – Caractéristiques des cas graves de grippe admis en réanimation pour les variables sexe, âge, classe d’âge, région, type d’hôpital et unité médicale pour chaque source, France entière, Semaine 40 de 2021 à Semaine 30 de 2022.....	<b>35</b>
<b>Tableau 3</b> - Caractéristiques des cas graves de grippe admis en réanimation pour les variables sexe, âge, classe d’âge, région, type d’hôpital et unité médicale selon les résultats du croisement, France entière, Semaine 40 de 2021 à Semaine 30 de 2022.....	<b>36</b>
<b>Tableau 4</b> – Comparaison de l’estimateur de Chapman avec le <i>Mark Recapture Ratio</i> pour les exhaustivités SGS et PMSI globales par région.....	<b>43</b>

## Liste des abréviations

<b>A(H1N1)<sub>pdm09</sub></b>	Virus A(H1N1) à l'origine de la pandémie de 2009
<b>AIC</b>	<i>Akaike Information Criterion</i> (Critère d'information d'Akaike)
<b>ARN</b>	Acide ribonucléique
<b>ARS</b>	Agence Régionale de Santé
<b>BCS</b>	<i>Best case scenario</i>
<b>BEH</b>	Bulletin épidémiologique hebdomadaire
<b>BIC</b>	<i>Bayesian Information Criterion</i> (Critère d'Information Bayésien)
<b>CDC</b>	<i>Centers for Disease Control and Prevention</i>
<b>CépiDc</b>	Centre d'épidémiologie sur les causes médicales de Décès
<b>CH</b>	Centre hospitalier
<b>CHU</b>	Centre hospitalier universitaire
<b>CIM-10</b>	Classification Internationale des Maladies (dixième version)
<b>CNIL</b>	Commission Nationale de l'Informatique et des Libertés
<b>CNR</b>	Centre National de Référence
<b>CR</b>	Cellule régionale (Santé publique France)
<b>DROM</b>	Départements et Régions d'Outre-Mer
<b>Ecsir</b>	Exhaustivité des cas sévères d'influenza en région
<b>EHPAD</b>	Établissement d'hébergement de personnes âgées dépendantes
<b>FINESS</b>	Fichier national des établissements sanitaires et sociaux
<b>GHM</b>	Groupes Homogènes de Malades
<b>HAS</b>	Haute Autorité de Santé
<b>IC</b>	Intervalle de confiance
<b>INSEE</b>	Institut National de la Statistique et des Études Économiques
<b>IRA</b>	Infection respiratoire aiguë
<b>MRR</b>	<i>Mark Recapture Ratio</i>
<b>OMS</b>	Organisation Mondiale de la Santé
<b>OR</b>	Odds Ratio
<b>OSCOUR®</b>	Organisation de la surveillance coordonnée des urgences
<b>PMSI-MCO</b>	Programme de Médicalisation des Systèmes d'Information (PMSI) en médecine, chirurgie, obstétrique et odontologie (MCO).
<b>RPU</b>	Résumé de passage aux urgences
<b>RSA</b>	Résumé de Sortie Anonyme
<b>RSS</b>	Résumé de Sortie Standardisé
<b>RUM</b>	Résumé d'Unité Médicale
<b>SGS</b>	Surveillance des Gripes Sévères
<b>SI-VIC</b>	Système d'information pour le suivi des victimes
<b>SNDS</b>	Système National des Données de Santé
<b>SurSaUD®</b>	Système de surveillance sanitaire des urgences et des décès
<b>T2A</b>	Tarifcation à l'activité
<b>UM</b>	Unité Médicale
<b>USC</b>	Unité de Soins Continus
<b>USI</b>	Unité de Soins Intensifs

## Résumé court

En France, le système de surveillance des cas graves de grippe admis en réanimation repose sur une déclaration de cas par des services de réanimation sentinelles. L'évaluation des systèmes de surveillance est essentielle pour mesurer leur performance et permettre leur amélioration. C'est pourquoi la surveillance des cas graves de grippe admis en réanimation a été évaluée de manière régulière depuis 2017 par le biais d'une analyse d'exhaustivité (études Ecsir).

La présente étude est la première étude Ecsir depuis le début de la pandémie de COVID-19 et porte sur la saison grippale 2021-2022. L'exhaustivité de la surveillance a été évaluée par croisement des données de la surveillance des grippes sévères (SGS) et de la base hospitalière PMSI en appliquant la méthode de capture-recapture à deux sources. Deux estimateurs, ceux de Chapman et de Chao, ont été employés et comparés. Les résultats ont également été présentés sous la forme d'un intervalle de plausibilité, noté ici  $I_{Ch-BCS}$ , afin de prendre en compte les biais potentiels inhérents à la méthode.

Au niveau national, l'exhaustivité SGS était vraisemblablement comprise dans l'intervalle  $I_{Ch-BCS}$  [12,3-25,5%] pour les analyses globales et [20,0-34,1%] pour les analyses restreintes aux services sentinelles. Ces valeurs étaient plutôt faibles par rapport aux résultats des études précédentes mais la comparabilité était limitée étant donné l'emploi d'une méthodologie différente pour l'obtention des estimations. L'exhaustivité PMSI était vraisemblablement comprise dans l'intervalle [44,7-93,2%] pour les analyses globales et [53,1-90,9%] pour les analyses restreintes aux services sentinelles. De fortes disparités ont été observées selon les régions. La surveillance SGS présente une meilleure exhaustivité pour les classes d'âge les plus âgées ainsi que pour les hôpitaux de type Centre Hospitalier (CH).

Il est probable que la surveillance des gripes sévères ait été fortement affectée par la pandémie de la COVID-19 (surcharge et moindre participation à la surveillance des services de réanimation, erreurs de diagnostic/codage), ce qui pourrait expliquer ces faibles exhaustivités. De manière générale, les résultats obtenus sont à prendre avec précaution mais peuvent donner une tendance générale et des ordres de grandeurs pris par les valeurs d'exhaustivité. Une description plus fine des modalités de recueil et transmission de données SGS et PMSI ainsi que l'investigation concernant des possibles dépendances entre les deux systèmes est recommandée. Une nouvelle sélection de services sentinelles privilégiant les services participant de manière active pourrait également être à considérer.

**Mots clés : surveillance, capture-recapture, exhaustivité, grippe, réanimation, France**

Suggestion de citation : Yalikun, I. (2023). *Analyse de l'exhaustivité de la surveillance des cas graves de grippe admis en réanimation, saison 2021-2022* (p. 64). Mémoire Master 2 ESMIHA. École nationale Vétérinaire d'Alfort - CIRAD

## Résumé long

### Contexte

La grippe est une infection respiratoire aiguë touchant des millions de personnes chaque année. D'après Santé publique France, 2 à 6 millions de personnes sont touchées par an par la grippe en France et environ 10 000 décès annuels sont attribuables à cette maladie, principalement chez les personnes âgées. La grippe est une infection virale contagieuse dont la transmission se réalise essentiellement par voie aérienne via des gouttelettes mais également par contact avec des surfaces contaminées. La maladie est caractérisée par de la fièvre, de la toux, une rhinorrhée (écoulement nasal), des céphalées (maux de tête), des myalgies et arthralgies (douleurs musculaires et articulaires) et de l'asthénie (malaise, fatigue). Le fardeau de la grippe, tant sanitaire, sociétal qu'économique justifie la mise en place d'une surveillance de cette maladie à l'échelle nationale. Cette surveillance, organisée par Santé publique France, se décline selon plusieurs modalités dont la surveillance des cas sévères de grippe admis en réanimation. Cette surveillance repose sur le signalement des cas par des services de réanimation sentinelles via des fiches transmises aux cellules régionales (CR) de Santé publique France.

Les performances d'un système de surveillance doivent être évaluées afin de permettre son amélioration et vérifier l'adéquation avec ses objectifs. Ainsi, en France, le système de surveillance des cas graves de grippe admis en réanimation est évalué par le biais des études d'Exhaustivité des cas sévères d'influenza en région (Ecsir). Initiées en 2017, les études Ecsir ont par la suite été menées de manière régulière. Leur objectif principal consiste à estimer l'exhaustivité de la surveillance des gripes sévères (SGS) via la méthode de capture-recapture à deux sources. La présente étude est la première étude Ecsir depuis le début de la pandémie et porte sur la saison grippale 2021-2022.

La méthode de capture-recapture permet d'estimer le nombre total d'individus dans une population. A partir du croisement de deux sources de données, les cas communs sont identifiés. Connaissant le nombre de cas communs ainsi que le nombre de cas de chaque source, il est possible d'estimer le nombre total d'individus d'une population. Cela inclut les individus qui n'ont été détectés par aucune source. L'exhaustivité d'une source est ensuite calculée en divisant le nombre de cas détectés par une source par le nombre de cas total estimé. Une grande valeur d'exhaustivité indique une bonne performance de la surveillance. De nombreux estimateurs permettent d'approximer le nombre de cas total. Dans le cadre d'une capture-recapture à deux sources, les plus courants sont les estimateurs dits de Lincoln-Petersen (ou *Mark Recapture Ratio* (MRR)), Chapman et Chao.

L'application de la méthode de capture-recapture répond à des conditions d'application assez strictes et qui peuvent varier selon les estimateurs. De manière générale, les conditions à vérifier sont les suivantes : tous les cas identifiés doivent être des vrais cas ; tous doivent être détectés dans les mêmes conditions spatio-temporelles ; l'application de la méthode doit porter sur une population close ; tous les vrais cas communs doivent être identifiés ; tous les cas de la population ont la même probabilité d'être identifiés par une source (homogénéité de capture) ; la probabilité d'être détecté par une source est indépendante de la probabilité d'être détecté par une autre source (indépendance des sources). Le non-respect des conditions d'application peut amener à des estimations biaisées, notamment pour l'estimateur de Chapman et le MRR. L'estimateur de Chao est moins sensible aux hypothèses d'homogénéité de capture et d'indépendance des sources. En présence d'hétérogénéité, l'estimateur de Chao donne un minorant du nombre de cas total, noté N.

### Matériel et méthodes

Les deux sources utilisées dans le cadre de notre étude sont les données issues du système que l'on cherche à évaluer (SGS) et les données provenant de la base hospitalière de séjours PMSI. Les

données SGS ont été obtenues par compilation des données transmises par les cellules régionales et permet l'obtention de la base nationale de cas de grippe sévère admis en réanimation pour la saison 2021-2022. Une base regroupant des services de réanimation ayant participé à la surveillance a également été mise à disposition dans l'objectif de la fusionner avec chacune des bases et d'attribuer à chaque cas des informations sur le service de réanimation correspondant. L'accès à la base PMSI a nécessité une demande d'autorisation d'extraction hors du Système National des Données de Santé (SNDS). Cette extraction a concerné les séjours hospitaliers avec un code diagnostic de grippe (code J09 à J11 selon la classification CIM-10) au sein d'une unité médicale (UM) de réanimation (UM 01, 02, 03), en diagnostic principal ou associé. L'accès et le traitement des données est permis par une autorisation permanente de la CNIL.

Les données des deux bases ont été restreintes entre la semaine 40 de 2021 et la semaine 20 de 2022. Pour la Guadeloupe et la Martinique, cela s'étendait jusqu'à la semaine 30 de 2022. Ces périodes correspondaient à la période de surveillance de la grippe pour 2021-2022. Les semaines considérées sont les semaines d'admission en réanimation. Les bases ont par la suite été nettoyées (par exemple suppression des doublons), certaines notations ont été harmonisées et des variables ont été rajoutées (par exemple la variable âge dans la base SGS). L'identification des cas communs entre les bases SGS et PMSI a été réalisée grâce à un croisement des deux bases sur 5 variables communes : le numéro FINESS géographique, le sexe, l'âge, le département de domicile et la date d'admission en réanimation. Cinq requêtes différentes ont été exécutées pour identifier les cas communs. Ces requêtes vérifiaient des égalités strictes ou autorisaient pour certaines des variations de valeur entre les deux bases. L'objectif était de rester spécifique tout en permettant l'identification d'un maximum de cas communs. La base de données finale comprend les cas communs et les cas qui ont été identifiés uniquement dans la base SGS ou le PMSI. Les cas communs PMSI sont ceux qui ont été gardés car les données PMSI sont considérées comme les plus fiables. Les cas ont été désignés comme étant sentinelle s'ils avaient été déclaré par un service initialement sentinelle ou par un service non initialement sentinelle mais ayant participé à la surveillance SGS.

Les conditions d'application étaient globalement vérifiées. Néanmoins, des incertitudes subsistaient pour la validité des hypothèses d'homogénéité de capture et d'indépendance des sources. Les estimateurs de Chapman et Chao ont tous deux été employés dans le but de pouvoir comparer leurs résultats. Ceux-ci étant plus ou moins sensibles à certaines de ces hypothèses, leur comparaison peut être pertinente pour tester la validité des conditions d'application. Étant donné les incertitudes liées au respect des conditions d'application et de la fiabilité des estimations, les résultats d'exhaustivité ont également été donnés sous la forme d'un intervalle dit de plausibilité, noté ici  $I_{Ch-BCS}$ , construit à partir de la borne inférieure de l'intervalle de confiance (IC) bootstrap 95% de l'estimateur de Chao pour borne inférieure et l'exhaustivité calculée dans une situation idéale (« Best case scenario », noté BCS) pour borne supérieure. Les IC ont été calculés selon la méthode du bootstrap paramétrique en considérant une distribution binomiale avec 10 000 ré-échantillonnages.

Une fois le nombre total de cas de grippe sévère admis en réanimation estimé, une analyse de l'exhaustivité à l'échelle nationale et métropolitaine a été menée. Des analyses stratifiées ont été réalisées selon les variables classe d'âge, région, type d'hôpital et période. Ces analyses ont également été réalisées en restreignant les analyses aux services sentinelles. Le taux d'incidence pour cent mille habitants a aussi été estimé en divisant le nombre de cas total estimé pour chaque région par les données de population régionales de 2022.

Toutes les étapes de nettoyage des bases de données ainsi que d'analyse ont été réalisées sous R 4.1.3.

## Résultats

Au total, la base SGS comportait 475 cas et la base PMSI comportait 1736 cas. Les requêtes de croisement ont identifié 349 cas communs.

Au niveau national et pour les analyses globales, l'exhaustivité de la surveillance SGS calculée à partir du nombre de cas estimé par la méthode de Chapman était de 20,1% (IC 95%, [18,2-22%]) et l'exhaustivité du PMSI était de 73,5% [66,8-80,5%]. L'estimateur de Chao donnait une valeur de 13,6% (IC 95%, [12,3-14,8%]) pour l'exhaustivité SGS et 49,6% (IC 95%, [44,7-54,3%]) pour l'exhaustivité PMSI. L'intervalle de plausibilité  $I_{Ch-BCS}$  était de [12,3-25,5%] pour l'exhaustivité SGS et de [44,7-93,2%] pour l'exhaustivité PMSI.

Au niveau des analyses restreintes aux services sentinelles, l'exhaustivité de la surveillance SGS était de 27,6% [25,2-30,1%] avec l'estimateur de Chapman et de 21,9% [20,0-23,8%] avec l'estimateur de Chao. Pour le PMSI, l'exhaustivité calculée selon les deux méthodes était, respectivement, 73,5% [67,0-80,0%] et 58,3% [53,1-63,5%]. L'intervalle  $I_{Ch-BCS}$  était de [20,0-34,1%] pour l'exhaustivité SGS et [53,1-90,9%] pour l'exhaustivité PMSI. Les données restreintes à la métropole donnaient des résultats assez similaires à ceux obtenus avec les analyses globales et restreintes aux services sentinelles.

Des disparités d'exhaustivité étaient observées entre les strates pour les analyses stratifiées par classe d'âge, région et type d'hôpital. En ce qui concerne la variable « période », l'exhaustivité était similaire pour les deux périodes considérées. Les valeurs d'exhaustivité issues des analyses restreintes aux services sentinelles étaient généralement plus élevées qu'avec les analyses globales. La surveillance SGS présentait une meilleure exhaustivité pour les classes d'âge les plus âgées ainsi que pour les hôpitaux de type Centre Hospitalier (CH).

Au niveau national, l'incidence des cas graves de grippe admis en réanimation était estimée à 3,5 [3,2-3,8] avec l'estimateur de Chapman et 5,2 [4,7-5,7] avec l'estimateur de Chao, avec de fortes disparités selon les régions.

## Discussion

L'étude Ecsir réalisée sur la saison grippale 2021-2022 est la première analyse d'exhaustivité du système de surveillance des gripes sévères depuis la pandémie de COVID-19. Cette dernière a eu un impact majeur sur la surveillance avec par exemple le passage en surveillance conjointe COVID/grippe et la saturation des services de réanimation. Une évaluation du système de surveillance était donc nécessaire pour en mesurer la performance après plusieurs années de pandémie COVID.

Les valeurs d'exhaustivité SGS estimées semblaient très faibles par rapport aux exhaustivités globales estimées de 47% [47-48%] pour la période 2009-2013, 46,1% [45,5-46,8%] pour la période 2017-2018 et 36,3% [35,6-37,0%] pour la période 2018-2019 (données non publiées). La comparabilité avec les résultats des études précédentes était cependant limitée étant donné le recours à une méthodologie différente. Une analyse des données 2021-2022 reprenant les mêmes modalités que la dernière étude Ecsir sera réalisée ultérieurement pour permettre la comparabilité. Les analyses restreintes au niveau métropolitain n'ont pas vraiment montré de différences par rapport aux analyses nationales étant donné la faible proportion de cas déclarés par les DROM. Les exhaustivités estimées au niveau régional étaient très disparates. Au niveau global, ceci pouvait témoigner de différences dans le nombre de services suivis selon les régions. Au niveau des analyses restreintes aux services sentinelles, ces résultats pourraient être expliqués par des taux de participation à la surveillance variables selon les régions. Le taux d'incidence estimé montrait que les régions les plus touchées par la grippe étaient Mayotte et l'Île-de-France et que la Bretagne était la région la moins touchée. Cette variabilité pouvait être multifactorielle (différences dans la circulation du virus, nombre de lits de réanimation disponibles par région etc.).

Alors que les analyses sur données globales permettent surtout d'estimer la proportion des cas signalés sur l'ensemble des cas graves de grippe en France, la performance du réseau de surveillance SGS est, elle, estimée grâce aux données sentinelles. Les résultats d'exhaustivité SGS étaient plus élevés lorsque l'analyse était restreinte aux données sentinelles, ce qui était attendu. Une variabilité importante a été observée selon les régions. La stratification par classe d'âge a également montré que le réseau de surveillance des grippes sévères présentait en moyenne une exhaustivité plus élevée pour la catégorie des 18-64 ans ainsi que les 75 ans et plus. L'exhaustivité SGS était également en moyenne meilleure pour les hôpitaux de type Centre Hospitalier (CH). Cela peut s'expliquer par le fait que de nombreux services ayant été retenus lors du passage en surveillance sentinelle étaient des services de CH.

Nos analyses ont mis en évidence un niveau de signalement des cas plutôt constant au cours du temps, avec une courbe épidémique quasi identique dans son allure par rapport à celle du PMSI. Les valeurs d'exhaustivité SGS restaient malgré tout plutôt faibles, même pour les données sentinelles. Nous pouvons supposer que la pandémie de la COVID-19 a eu un impact important sur le signalement des cas graves de grippe admis en réanimation. Une description détaillée du parcours de recueil de données et de signalement des deux bases serait un atout non négligeable pour l'interprétation des données à l'avenir.

Globalement, dans nos analyses, les valeurs du nombre de cas total estimées par l'estimateur de Chao étaient supérieures de façon significative par rapport aux estimations de Chapman (non recoupement des IC 95%). La présence d'un lien de dépendance positive entre les sources ou une hétérogénéité de capture pourrait expliquer les différences observées entre les deux estimations. Nos résultats ont montré des disparités dans la distribution des effectifs par région et par classe d'âge. L'observation de valeurs de N estimées par stratification significativement différente du N estimé globalement pourrait être en faveur de la présence d'une hétérogénéité de capture. La valeur prise par l'estimateur de Chao pour certaines régions se rapprochait de celle de l'estimateur de Chapman pour les deux bases, de même pour les estimations de N par stratification dans le cadre des analyses restreintes aux services sentinelles. Nous pouvons alors supposer que ces situations se rapprocheraient davantage des conditions où les conditions d'homogénéité et d'indépendance sont vérifiées. Il est possible que cela soit principalement dû au fait que la restriction des analyses aux services sentinelles supprime un facteur d'hétérogénéité important, à savoir le statut sentinelle d'un service. Les différences dans les modalités d'animation du réseau de surveillance selon les régions pourraient également impacter l'exhaustivité régionale.

Une description plus détaillée des modalités de recueil des données pour les bases SGS et PMSI ainsi que des modalités d'animation par région pourrait être hautement profitable pour les prochaines études Ecsir

Des incertitudes subsistaient encore quant à la vérification d'une hypothèse propre à l'estimateur de Chao (probabilité de capture d'un individu constante entre les sources) ainsi que la sélection des cas issus des services sentinelles. Si tous ces éléments peuvent nous faire douter de la précision ou de la fiabilité des résultats et incitent à la prudence quant à l'interprétation des estimations brutes, les données présentées ici sont néanmoins importantes pour suivre les tendances d'une année à l'autre et identifier des leviers d'amélioration. Ces informations sont utiles pour l'évaluation de la surveillance, ce qui est en faveur d'une analyse d'exhaustivité régulière par le biais des études Ecsir.

# Table des matières

REMERCIEMENTS .....	2
LISTE DES FIGURES ET DES TABLEAUX.....	3
LISTE DES ABREVIATIONS.....	4
RESUME COURT .....	5
RESUME LONG .....	6
INTRODUCTION.....	12
<b>I - SYNTHÈSE BIBLIOGRAPHIQUE .....</b>	<b>13</b>
<b>1. LA GRIPPE : UNE INFECTION RESPIRATOIRE AU LOURD FARDEAU .....</b>	<b>13</b>
<b>2. LA SURVEILLANCE DE LA GRIPPE EN FRANCE .....</b>	<b>14</b>
2.1 <i>Surveillance de la grippe en médecine de ville et en collectivités de personnes âgées .....</i>	<i>14</i>
2.2 <i>Surveillance de la grippe dans le milieu hospitalier.....</i>	<i>15</i>
2.3 <i>Surveillance de la grippe en population générale.....</i>	<i>15</i>
2.4 <i>Surveillance de la mortalité attribuable à la grippe .....</i>	<i>15</i>
<b>3. FOCUS SUR LA SURVEILLANCE DES CAS GRAVES DE GRIPPE ADMIS EN REANIMATION ET L'ÉVALUATION DE SON EXHAUSTIVITÉ .....</b>	<b>16</b>
3.1 <i>Origine et modalités de la surveillance des gripes sévères (SGS) .....</i>	<i>16</i>
3.2 <i>L'étude Ecsir (Exhaustivité des cas sévères d'influenza en région).....</i>	<i>17</i>
3.3 <i>Impact de la pandémie du COVID-19 sur la surveillance des gripes sévères et sur l'analyse de son exhaustivité .....</i>	<i>17</i>
<b>4. ANALYSE DE L'EXHAUSTIVITÉ DE LA SURVEILLANCE PAR LA MÉTHODE DE CAPTURE-RECAPTURE.....</b>	<b>18</b>
4.1 <i>Origine de la méthode de capture-recapture et application à l'épidémiologie.....</i>	<i>18</i>
4.2 <i>Conditions d'application de la méthode de capture-recapture .....</i>	<i>19</i>
4.3 <i>Estimateurs utilisés dans le cadre de la capture-recapture à deux sources .....</i>	<i>19</i>
4.4 <i>Sources de données utilisées pour l'analyse de l'exhaustivité de la surveillance des cas sévères de grippe admis en réanimation par capture-recapture .....</i>	<i>22</i>
<b>II – TRAVAIL PERSONNEL .....</b>	<b>23</b>
<b>1. RAPPEL DU CONTEXTE, PROBLÉMATISATION ET OBJECTIFS.....</b>	<b>23</b>
<b>2. MATÉRIEL ET MÉTHODES .....</b>	<b>23</b>
2.1 <i>Accès et description des bases de données SGS et PMSI .....</i>	<i>23</i>
2.2 <i>Traitement des bases de données SGS et PMSI en amont de l'analyse.....</i>	<i>24</i>
2.3 <i>Identification des cas communs par croisement des bases SGS et PMSI dans le cadre de la méthode de capture-recapture.....</i>	<i>27</i>
2.4 <i>Calcul de l'exhaustivité par la méthode de capture-recapture.....</i>	<i>29</i>
2.5 <i>Analyses réalisées .....</i>	<i>32</i>
<b>3. RESULTATS .....</b>	<b>33</b>
3.1 <i>Description des bases SGS et PMSI .....</i>	<i>33</i>
3.2 <i>Croisement des bases SGS et PMSI .....</i>	<i>36</i>
3.3 <i>Estimation de l'exhaustivité des bases SGS et PMSI par capture-recapture .....</i>	<i>39</i>
<b>4. DISCUSSION.....</b>	<b>46</b>
<b>III - CONCLUSION.....</b>	<b>51</b>
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>52</b>
<b>ANNEXES.....</b>	<b>56</b>
ANNEXE 1 – FICHE DE SIGNALEMENT DES CAS SEVERES DE GRIPPE ADMIS EN REANIMATION DANS LE CADRE DE LA SURVEILLANCE SGS.....	56

ANNEXE 2 – DISTRIBUTION DES EFFECTIFS SELON PLUSIEURS VARIABLES COMMUNES AUX BASES SGS ET PMSI POUR LES ANALYSES RESTREINTES AUX SERVICES SENTINELLES .....	59
ANNEXE 3 – NOMBRE DE CAS RECUPERES PAR CHAQUE REQUETE DE CROISEMENT .....	60
ANNEXE 4 – ESTIMATION DE L'EXHAUSTIVITE PAR SOURCE ET PAR PERIODE DE TEMPS POUR LES ANALYSES GLOBALES ET RESTREINTES AUX SERVICES SENTINELLES, FRANCE ENTIERE, SEMAINE 40 DE 2021 A SEMAINE 30 DE 2022 .....	60
ANNEXE 5 – ESTIMATION DE L'EXHAUSTIVITE PAR SOURCE ET PAR TYPE D'HOPITAL POUR LES ANALYSES GLOBALES ET RESTREINTES AUX SERVICES SENTINELLES, FRANCE ENTIERE, SEMAINE 40 DE 2021 A SEMAINE 30 DE 2022 .....	61
ANNEXE 6 – CARTES DES EXHAUSTIVITES MOYENNES SGS ET PMSI POUR LES ANALYSES RESTREINTES AUX SERVICES SENTINELLES .....	62
ANNEXE 7 – ESTIMATION DU NOMBRE TOTAL DE CAS DE GRIPPE SEVERES ADMIS EN REANIMATION PAR STRATIFICATION SELON LES ESTIMATEURS DE CHAPMAN ET CHAO .....	64
ANNEXE 8 – EXPRESSION DE L'ESTIMATEUR DE CHAO CORRIGEE POUR LES FAIBLES EFFECTIFS .....	64

## Introduction

La grippe est une infection respiratoire aiguë touchant des millions de personnes chaque année. D'après Santé publique France, 2 à 6 millions de personnes sont touchées par an par la grippe en France et environ 10 000 décès annuels sont attribuables à cette maladie, principalement chez les personnes âgées. Une surveillance de la grippe a été mise en place au niveau national et à plusieurs échelles allant de la surveillance en médecine de ville à la surveillance des collectivités de personnes âgées, tout en passant par la surveillance en milieu hospitalier. Chacune se décline selon plusieurs dispositifs ayant chacun des populations d'intérêt et des objectifs spécifiques. On retrouve par exemple en milieu hospitalier une surveillance des cas graves de grippe admis en réanimation. Son objectif consiste en particulier en la description des cas graves de grippe et en l'estimation de la sévérité d'une épidémie.

L'évaluation des systèmes de surveillance est essentielle pour en mesurer les performances, l'adéquation avec les objectifs fixés au départ ainsi qu'en permettre l'amélioration. Celle-ci peut être quantitative ou qualitative. Le système de surveillance des cas graves de grippe admis en réanimation est régulièrement évalué de manière quantitative dans le cadre des études Ecsir depuis 2017. L'objectif principal de ces études est d'estimer l'exhaustivité de la surveillance des cas graves de grippe grâce à l'application de la méthode de capture-recapture. Les résultats de ces études permettent notamment de mieux apprécier la représentativité des chiffres d'incidence de la maladie calculée chaque année.

A partir de la saison 2018-2019, la surveillance des gripes sévères a évolué d'une surveillance exhaustive à une surveillance sentinelle. Aucune étude Ecsir n'a été réalisée depuis la pandémie de COVID-19. Cette dernière a eu un impact important sur la surveillance de la grippe, la circulation de des virus grippaux ainsi que sur l'activité hospitalière. Une évaluation du système de surveillance des cas de grippe sévère admis en réanimation s'avère donc nécessaire afin d'évaluer les performances de cette surveillance après plusieurs années de pandémie. Ainsi, l'objectif principal de la présente étude est de réaliser une étude d'exhaustivité pour cette surveillance. L'étude Ecsir que nous avons réalisée est la première depuis le début de la pandémie et porte sur la saison grippale 2021-2022.

Tout d'abord, une synthèse des connaissances sur la grippe et sa surveillance en France sera présentée. Le principe des études Ecsir ainsi que la méthode de la capture-recapture employée dans ce cadre seront détaillées. Par la suite, le travail d'analyse mené sera décrit, l'objectif principal étant de présenter les résultats d'exhaustivité de la surveillance des cas graves admis en réanimation pour la saison grippale 2021-2022. Enfin, les résultats seront discutés et interprétés.

# I - Synthèse bibliographique

## 1. La grippe : une infection respiratoire au lourd fardeau

La grippe est une infection respiratoire aiguë (IRA) causée par les Influenzavirus, des virus à ARN monocaténaire à polarité négative appartenant à la famille des *Orthomyxoviridae*. Ils sont distingués en quatre types nommés A, B, C et D. Seuls les virus de type A, B et C peuvent infecter l'espèce humaine. Les virus de type A et B sont responsables d'épidémies et certains virus de type A sont à l'origine de pandémies, comme par exemple la pandémie de 2009 due au virus de la grippe A(H1N1)<sub>pdm09</sub> (OMS ; CDC, c ; Uyeki *et al.*, 2022)

La maladie est caractérisée par de la fièvre, de la toux, une rhinorrhée (écoulement nasal), des céphalées (maux de tête), des myalgies et arthralgies (douleurs musculaires et articulaires) et de l'asthénie (malaise, fatigue) (OMS). La grippe est une infection virale contagieuse circulant principalement sur une période s'étendant d'octobre à avril dans les régions tempérées et pendant toute l'année dans les régions tropicales (Viboud *et al.*, 2006 ; Uyeki *et al.*, 2022 ; OMS). La transmission se réalise essentiellement par voie aérienne via des gouttelettes mais également par contact avec des surfaces contaminées. Même si la majorité des sujets se rétablissent généralement en une semaine, les symptômes peuvent parfois durer jusqu'à deux semaines ou plus et dans certains cas mener à un décès. Certaines catégories de population sont plus à risque de formes graves, notamment les personnes âgées, les femmes enceintes, les jeunes enfants et les personnes immunodéprimées (Uyeki *et al.*, 2022 ; OMS ; CDC, b).

Les traitements visent généralement à diminuer l'intensité des symptômes. Dans certains cas, des médicaments antiviraux peuvent être prescrits, notamment pour des patients à risque. La prévention contre la grippe est principalement menée par des campagnes de vaccination ainsi que la promotion des gestes barrières. (Uyeki *et al.*, 2022 ; OMS ; Santé publique France, a ; CDC, e).

A l'échelle mondiale, les épidémies de grippe saisonnière sont responsables d'environ 3 à 5 millions de cas graves et de 290 000 à 650 000 décès chaque année (OMS). En France, 2 à 6 millions de personnes sont touchées par la grippe chaque année et environ 10 000 décès annuels sont attribuables à cette maladie, dont 90% sont des personnes âgées de plus de 65 ans (Santé publique France, a). Le fardeau des gripes pandémiques est tout aussi important voire supérieur. On estime par exemple que la grippe pandémique A(H1N1) de 2009 a causé entre 151 700 et 575 400 dans le monde lors de sa première année de circulation et celle de 1918 au moins 50 millions (Johnson et Mueller, 2002 ; Dawood *et al.*, 2012 ; CDC, a). Ces épidémies peuvent être associées à un important absentéisme au travail et donc à une perte de productivité ainsi qu'à une forte pression sur les structures hospitalières (OMS).

En outre, les virus Influenza présentent une forte capacité à évoluer génétiquement. Les deux principaux mécanismes d'évolution sont le glissement antigénique ou *drift* qui consiste en des modifications des séquences des protéines de surface d'hémagglutinine et de neuraminidase, pouvant notamment permettre au virus d'échapper à la réponse immunitaire ; et la cassure antigénique ou *shift* qui consiste en un réassortiment du matériel génétique de plusieurs virus amenant à la création d'un nouveau virus auquel la population présente une faible immunité. Le phénomène de *drift* est généralement à l'origine de la circulation de nouveaux variants tandis que le *shift* contribue à l'émergence de nouveaux virus pouvant être à l'origine de pandémies (Webster *et al.*, 2014 ; Krammer *et al.*, 2018 ; CDC, d). Cette capacité des Influenzavirus à évoluer explique la nécessité de produire de nouveaux vaccins adaptés aux souches circulantes chaque année ainsi que l'organisation de campagnes annuelles de vaccination (OMS ; Uyeki *et al.*, 2022). Des Influenzavirus circulent également chez les espèces animales, le principal réservoir étant constitué par les oiseaux sauvages. Certaines souches de grippe sont zoonotiques, c'est-à-dire qu'elles peuvent se transmettre d'humains à animaux et inversement (Krammer *et al.*, 2018). Des réassortiments peuvent alors se produire entre souches

d'Influenzavirus provenant d'espèces différentes. Par exemple, le virus pandémique A(H1N1)<sub>pdm09</sub> a été décrit comme provenant d'un triple réassortiment entre des gènes de virus grippaux porcins, aviaires et humains (Dawood *et al.*, 2009).

Ainsi, le fardeau sanitaire, sociétal et économique de la grippe, causée par des virus en constante évolution, explique la nécessité d'un suivi continu et par conséquent la mise en place d'une surveillance de la grippe.

## 2. La surveillance de la grippe en France

En France, de nombreux dispositifs contribuent à la surveillance de la grippe. Cette surveillance, coordonnée par Santé publique France, a pour objectif de permettre la détection d'une épidémie de grippe, réaliser une analyse descriptive de l'épidémie (description spatio-temporelle du phénomène et des populations concernées), permettre d'identifier et de suivre les souches circulantes et évaluer l'efficacité des mesures de contrôle mises en place pour endiguer l'épidémie, notamment la vaccination (Santé publique France, b).

### 2.1 Surveillance de la grippe en médecine de ville et en collectivités de personnes âgées

Au niveau de la médecine de ville, trois dispositifs peuvent être distingués : la surveillance par le réseau Sentinelles et la surveillance par le réseau SOS médecins qui assurent une surveillance de type syndromique (surveillance automatisée d'indicateurs de santé) et la surveillance par le Centre National de Référence (CNR) des virus des infections respiratoires qui assure une surveillance virologique (Santé publique France, f).

Le réseau Sentinelles regroupe un ensemble de 1252 médecins généralistes et 138 pédiatres volontaires (chiffres au 1<sup>er</sup> janvier 2022) qui signalent l'observation de dix indicateurs de santé dont les syndromes grippaux. La définition de cas utilisée est l'apparition soudaine d'une fièvre supérieure à 39°C avec myalgies et signes respiratoires. Le nombre de consultations pour syndrome grippal est suivi de manière hebdomadaire et permet l'estimation de l'incidence des syndromes grippaux. La définition de cas a cependant évolué suite à la pandémie de la COVID-19, amenant à surveiller plus généralement les IRA à la place des syndromes grippaux dans le but de pouvoir inclure la COVID-19 dans la surveillance. La définition de cas correspond alors à toute consultation pour une IRA définie par une fièvre apparaissant de manière brutale (ou sensation de fièvre) et des signes respiratoires. Les données sont transmises de manière continue au serveur du réseau Sentinelles avant d'être analysées et publiées sur le site Sentiweb chaque semaine (Réseau Sentinelles ; Santé publique France, f)

Des prélèvements nasopharyngés peuvent être réalisés par les acteurs du réseau Sentinelles sur un échantillon représentatif de patients présentant un syndrome grippal puis analysés par le CNR des virus des infections respiratoires. Cela permet l'identification des virus grippaux circulants et contribue à l'adaptation des vaccins antigrippaux aux virus détectés, essentielle pour assurer leur efficacité. Suite à la pandémie de la COVID-19, les prélèvements salivaires ont été préférés par rapport aux prélèvements nasopharyngés dans le but de pouvoir réaliser davantage de prélèvements (Institut Pasteur ; Santé publique France, f ; Équipes de surveillance de la grippe, 2022).

La surveillance des syndromes grippaux est également assurée par les professionnels de santé des associations SOS Médecins. La définition de cas utilisée est une fièvre soudaine supérieure à 38,5°C associée à des myalgies et des signes respiratoires. Les données anonymisées correspondant aux actes médicaux réalisés sont transmises de manière quotidienne et automatisée à Santé publique France via le système de surveillance sanitaire des urgences et des décès SurSaUD®. Cela permet une obtention automatisée et régulière d'indicateurs épidémiologiques comme par exemple la proportion de patients avec un diagnostic de grippe ou syndrome grippal sur l'ensemble des consultations SOS Médecins (Santé publique France, e ; Santé publique France, f).

Une surveillance des IRA est également mise en place dans les collectivités de personnes âgées, comme par exemple les Établissements d’hébergement de personnes âgées dépendantes (EHPAD), pour suivre les cas groupés. Ces derniers sont définis par l’apparition au minimum de 5 cas d’IRA chez les résidents dans un délai de 4 jours. Les signalements sont réalisés par le biais d’une plateforme accessible sur le site du ministère de la Santé (Santé publique France, f).

### **2.3 Surveillance de la grippe dans le milieu hospitalier**

Au niveau des hôpitaux, une surveillance syndromique est réalisée par le réseau OSCOUR® (Organisation de la surveillance coordonnées des urgences) qui fait également partie intégrante de SurSaUD®. Cette surveillance consiste en une extraction et une transmission automatisée des données du résumé de passage aux urgences (RPU). Environ 93% des passages aux urgences en France sont couverts par le réseau (chiffres de 2021). Cela permet de produire quotidiennement et automatiquement des indicateurs épidémiologiques à partir des données de consultation aux urgences ou d’une hospitalisation due à une grippe. L’indicateur principal est la proportion d’hospitalisations avec un diagnostic de grippe ou syndrome grippal parmi l’ensemble des hospitalisations, informant sur l’impact des épidémies de grippe à l’hôpital. Le codage du diagnostic suit la Classification Internationale des Maladies (CIM-10) (Santé publique France, d ; Santé publique France, f). Les codes utilisés pour la grippe sont les codes J09, J10 et J11.

Une surveillance des cas graves de grippe admis en réanimation est également menée au niveau hospitalier. Elle s’appuie sur un réseau de services de réanimation sentinelle qui transmettent les données sur les cas graves à Santé publique France via des fiches de signalement (Annexe 1). L’objectif est de suivre l’évolution de la dynamique épidémiologique des cas graves de grippe, évaluer la sévérité, réaliser une description de ces cas graves et utiliser les données pour permettre une meilleure organisation des services de réanimation (Santé publique France, f). Pour la saison grippale 2021-2022, 211 services de réanimation ont été sollicités dans le cadre de cette surveillance (Équipes de surveillance de la grippe, 2022).

Des données agrégées sont partagées entre certains laboratoires hospitaliers et le CNR des virus des infections respiratoires dans le cadre du réseau RELAB. Celui-ci est constitué d’une quarantaine de laboratoires de Centres Hospitaliers (CH) et Centres Hospitaliers Universitaires (CHU) principalement. Les données partagées sont par exemple le nombre de tests grippe réalisés et le nombre de gripes détectées ainsi que leur type. Ces données contribuent à la surveillance virologique. (Équipes de surveillance de la grippe, 2022).

### **2.3 Surveillance de la grippe en population générale**

Une surveillance en population générale existe également via la plateforme GrippeNet.fr dont l’objectif est de recueillir des données sur la population de France métropolitaine, en dehors du cadre hospitalier ou de la médecine de ville, grâce à des participants volontaires rapportant de manière hebdomadaire leurs symptômes. Cela permet notamment d’avoir accès à des données pour les personnes atteintes de grippe qui ne sont pas détectées par les autres moyens de surveillance comme par exemple ceux qui ne consultent pas pour leurs symptômes. Ce dispositif est coordonné par le réseau Sentinelles (Santé publique France, f ; GrippeNet).

### **2.4 Surveillance de la mortalité attribuable à la grippe**

Une surveillance syndromique des décès toute cause à partir des données de l’Institut National de la Statistique et des Études Économiques (INSEE) est réalisée dans le cadre du dispositif SurSaUD®. Elle permet de détecter des excès de mortalité, c’est-à-dire des différences significatives entre la

mortalité attendue (estimée à partir de données historiques) et celle observée chaque semaine. Environ 80% des décès à l'échelle du pays sont couverts par ce dispositif. Cependant, aucun lien ne peut être clairement établi avec des pathologies en l'absence de données sur les causes de décès.

L'estimation de la mortalité attribuable à la grippe nécessite donc une surveillance plus spécifique. Cette surveillance s'appuie sur les données de décès recensés parmi des cas groupés d'IRA en collectivité et les cas graves admis en réanimation ainsi que par l'utilisation d'un modèle statistique développé par Santé publique France estimant la mortalité attribuable à la grippe pour une épidémie à partir des données de mortalité toutes causes.

Les certificats électroniques de décès où la cause du décès a été attribuée à la grippe peuvent également permettre l'estimation de la mortalité attribuable à cette maladie. Ces certificats sont transmis au Centre d'épidémiologie sur les causes médicales de Décès (CépiDc), responsable de la production de statistiques nationales des causes de décès. Cependant, ce dispositif ne permet pas encore d'atteindre une bonne exhaustivité (couverture de 37% des décès mi-2022) et les délais d'obtention des données (environ 2 ans) ne permettent pas une utilisation en routine (Équipes de surveillance de la grippe, 2022 ; Santé publique France, 2023 ; Santé publique France, f ; CépiDc).

Toutes ces informations provenant des différents dispositifs de surveillance de la grippe sont par la suite analysées, interprétées et synthétisées par Santé publique France afin d'être publiées chaque semaine au cours de la saison hivernale dans un bulletin épidémiologique national hebdomadaire accessible en ligne. Un bilan épidémique est également publié chaque année dans le Bulletin épidémiologique hebdomadaire (BEH).

### **3. Focus sur la surveillance des cas graves de grippe admis en réanimation et l'évaluation de son exhaustivité**

#### **3.1 Origine et modalités de la surveillance des gripes sévères (SGS)**

Le sujet de cette étude porte plus spécifiquement sur la surveillance des cas graves de grippe admis en réanimation, aussi appelée surveillance des gripes sévères (SGS). Mise en place pour la surveillance des cas sévères du virus grippal pandémique A(H1N1)<sub>pdm09</sub> en 2009, elle s'est étendue aux cas graves de grippe A dès la saison grippale 2010-2011 avant de s'étendre à tous les cas graves de grippe à partir de la saison 2012-2013 (CNIL, 2017 ; Protocole Ecsir 2018, 2018). Cette surveillance a d'abord été à visée exhaustive avec l'objectif de mobiliser tous les services de réanimation avant de s'appuyer à partir de la saison 2018-2019 sur un réseau de 194 services de réanimation sentinelles en métropole (Santé publique France, c). Ce changement s'est justifié par le fait que la grande majorité des cas de grippe sévères se concentrent dans un nombre restreint de services de réanimation. Cela facilite également l'animation du réseau de surveillance (Loury *et al.*, 2017). La surveillance est néanmoins restée exhaustive dans les Départements et Régions d'Outre-Mer (DROM) où la dynamique de circulation de la maladie y est différente (climat tropical, pas de saisonnalité de la grippe).

L'objectif premier de cette surveillance lors de sa mise en place était de déterminer les caractéristiques des patients développant des formes graves de l'infection par le virus A(H1N1)<sub>pdm09</sub> à l'origine de la pandémie. Il avait notamment permis de mettre en évidence une prédominance des formes graves chez les personnes d'âge intermédiaire. La surveillance des cas graves de grippe admis en réanimation s'est par la suite poursuivie chaque année et se poursuit encore actuellement. Elle est coordonnée dans chaque région par les cellules régionales (CR) de Santé publique France et s'étend sur la période d'octobre à avril en métropole (généralement entre les semaines 45 et 15). Le rôle des CR est d'animer le réseau de services de réanimation participant à la surveillance et de centraliser les fiches de signalement des cas sévères de grippe transmises par ces services afin de procéder à la saisie

de ces données. La transmission des fiches est réalisée via une messagerie sécurisée de santé, une page de dépôt sécurisée ou par fax (Loury *et al.*, 2017 ; Santé publique France, c). La surveillance est généralement arrêtée en dehors de cette période en métropole étant donné que les cas sévères de grippe admis en réanimation hors du cadre temporel défini habituellement sont en général sporadiques.

### **3.2 L'étude Ecsir (Exhaustivité des cas sévères d'influenza en région)**

L'évaluation des systèmes de surveillance est un processus clé pour vérifier et améliorer le fonctionnement de ces systèmes. Jusqu'en 2017, aucune évaluation du système de surveillance des gripes sévères n'avait encore été réalisée, notamment en ce qui concerne la qualité des données et l'exhaustivité de la surveillance. C'est ainsi que l'étude d'Exhaustivité des cas sévères d'influenza en région (Ecsir) a été initiée pour pallier ce manque.

L'objectif principal de cette étude, tel que défini par Loury *et al.* (2017) dans le premier rapport Ecsir couvrant la période 2009-2013, est d'estimer le nombre de cas total de grippe sévère admis en réanimation et ainsi pouvoir évaluer l'exhaustivité de la surveillance des gripes sévères. C'est une évaluation de type quantitative. Les objectifs secondaires définis dans ce même rapport sont l'analyse des facteurs associés à des variations dans le temps et dans l'espace de l'exhaustivité, l'étude de l'impact de ces facteurs sur les taux d'incidence observés et la formulation d'améliorations de la surveillance et de l'interprétation des données recueillies. Cette première étude recommande une analyse régulière de l'exhaustivité, le but étant de pouvoir mieux interpréter les estimations d'incidence en les nuancant selon les résultats d'exhaustivité obtenus. Cela permet en outre d'avoir des informations sur la participation des services du réseau pour pouvoir en améliorer l'animation. Des études Ecsir ont par la suite été menées sur les saisons grippales 2017-2018 (Pivette *et al.*, 2019) et 2018-2019 (données non publiées).

### **3.3 Impact de la pandémie du COVID-19 sur la surveillance des gripes sévères et sur l'analyse de son exhaustivité**

La pandémie de la COVID-19 ayant affecté le monde à partir de début 2020 a eu un impact important sur l'organisation de la surveillance de la grippe et notamment celle des cas sévères. En mars 2020, la surveillance des cas de gripes sévères admis en réanimation est interrompue de manière prématurée pour devenir une surveillance des cas graves de COVID-19 admis en réanimation via le dispositif SI-VIC (Système d'information pour le suivi des victimes). Santé publique France met par la suite en place à partir d'octobre 2020 une surveillance conjointe des cas graves de grippe et de COVID-19, dans un contexte marqué notamment par des incertitudes sur les interactions entre le virus du SARS-CoV-2 (à l'origine de la maladie de la COVID-19) et d'autres virus respiratoires, justifiant cette surveillance conjointe (Santé publique France, c). La définition de cas correspond alors à tout « patient présentant une confirmation biologique de grippe ou d'infection à SARS-CoV-2 nécessitant une prise en charge en réanimation du fait de l'infection ou du fait de ses conséquences (décompensation de pathologie sous-jacente, surinfection bactérienne) – hors cas de syndrome inflammatoire multi-systémique pédiatrique lié à SARS-CoV-2 (Paediatric Inflammatory multisystem syndrome – PIMS) » (Annexe 1).

En outre, la circulation du SARS-CoV-2 et les mesures mises en place pour endiguer la pandémie ont eu un impact non négligeable sur la dynamique épidémiologique de la grippe. Par ailleurs, dans le bilan de la saison grippale 2021-2022 publié par Santé publique France, la comparaison de l'impact de l'épidémie 2021-2022 s'est basée sur les épidémies précédant la saison 2019-2020. Cette dernière avait été perturbée par la pandémie de COVID-19 et il n'y avait pas eu d'épidémie de grippe lors de la saison 2020-2021 (Équipes de surveillance de la grippe, 2022).

Il est également à noter qu'une analyse similaire à celle réalisée dans le cadre d'Ecsir a été menée pour la surveillance des admissions en réanimation pour la Covid-19 pour la période 2020-2021, témoignant une volonté d'un recours plus généralisé de ce type d'évaluation de la surveillance.

#### **4. Analyse de l'exhaustivité de la surveillance par la méthode de capture-recapture**

L'analyse de l'exhaustivité menée dans les différentes études Ecsir repose sur l'application de la méthode de capture-recapture. Cette méthode permet l'estimation du nombre total de cas graves de grippe admis en réanimation à partir du croisement d'au moins deux sources de données. Ce nombre total de cas estimé inclut à la fois les cas détectés par les systèmes de surveillances considérés et les cas non détectés (estimation). Le calcul de l'exhaustivité des sources de données servant au croisement est ensuite réalisé à partir du rapport du nombre de cas effectivement détectés par la source en question sur le nombre de cas total estimé.

##### **4.1 Origine de la méthode de capture-recapture et application à l'épidémiologie**

La méthode de capture-recapture a tout d'abord été appliquée dans le domaine de la zoologie et de la démographie pour estimer des effectifs de population animale ou humaine (IWGDMF, 1995 ; Gallay *et al.*, 2002). Les premières applications en démographie remonteraient, d'après Seber, au XVII-XVIIIème siècle avec par exemple son utilisation pour l'estimation de la taille de la population en France par Laplace en 1786 (Chao *et al.*, 2001). En 1894, Petersen applique cette méthode pour estimer l'effectif de poissons dans un fjord au Danemark et popularise la méthode. En 1949, Sekar et Deming ont eu recours à cette méthode pour estimer les taux de naissance et de décès. Leur travail est considéré comme pionnier dans l'application de la capture-recapture sur des populations humaines (IWGDMF, 1995).

Le principe de la méthode pour les populations animales repose sur le fait de réaliser une capture d'un échantillon de la population qu'on cherche à dénombrer. Par la suite, ces individus sont marqués et relâchés. Une seconde capture est ensuite réalisée. Il s'agit alors de dénombrer les individus précédemment marqués lors de la première capture au sein de ce second échantillon. En connaissant le nombre d'individus dans chacun des échantillons ainsi que le nombre d'individus communs aux échantillons, il devient alors possible d'estimer le nombre total d'individus (Chao *et al.*, 2001)

L'application de cette méthode en épidémiologie est récente, les premiers travaux datant de la seconde moitié du XXème siècle, notamment les études menées par Wittes et associés (Wittes et Sidel, 1968 ; Wittes *et al.*, 1974 ; IWGDMF, 1995). Le principe reste similaire à celui décrit précédemment : il s'agit de disposer d'au moins deux sources de données répertoriant les individus atteints par le phénomène ou la pathologie étudiée, de déterminer les cas communs aux sources considérées et par la suite estimer le nombre total d'individus de la population d'intérêt.

L'utilisation de la méthode de capture-recapture pour réaliser des analyses d'exhaustivité des systèmes de surveillances est une des spécialités de Santé publique France avec par exemple des études menées sur la surveillance des infections à méningocoques (Hubert et Desenclos, 1993), le SIDA (Spaccaferri *et al.*, 2010), la dengue (La Ruche *et al.*, 2013) ou encore la légionellose (Campese *et al.*, 2013).

## 4.2 Conditions d'application de la méthode de capture-recapture

La validité de l'application de cette méthode nécessite de vérifier certaines conditions, comme détaillé par Gallay *et al.* (2002). Dans un contexte d'application à l'épidémiologie (dénombrement de cas et non d'espèces animales), ces conditions peuvent être formulées comme ce qui suit :

- Tous les individus identifiés par les sources de données doivent être de vrais cas. La présence de faux cas aurait pour conséquence une surestimation du nombre total de cas et donc une sous-estimation de l'exhaustivité. Cela signifie généralement que les deux sources utilisent la même définition de cas.
- Les cas présents dans nos sources de données doivent être déTECTÉS dans les mêmes conditions spatio-temporelles. En effet, des périodes ou des zones géographiques d'étude différentes entre les sources peuvent limiter le nombre de cas communs lors du croisement de ces sources et être à l'origine d'une surestimation du nombre total de cas.
- Il ne doit pas y avoir de mouvements dans la population étudiée : on parle de population close. Comme pour le point précédent, l'étude sur une population ouverte peut amener à diminuer le nombre de cas communs et donc surestimer le nombre total de cas.
- Tous les vrais cas communs doivent être identifiés. Une surestimation ou une sous-estimation impacte, respectivement négativement ou positivement le nombre de cas total.
- Il doit y avoir homogénéité de la capture des cas. Cela signifie que tous les cas de la population d'étude ont la même probabilité d'être détectés par une source donnée. Cette probabilité peut varier selon les sources.
- Il y a indépendance des sources : la probabilité d'être détecté par une source ne doit pas être dépendante de la probabilité d'être détecté par une autre source. Lorsqu'il y a une dépendance positive, c'est-à-dire que la probabilité d'être détecté par une source augmente la probabilité d'être détecté dans une autre source, cela augmente le nombre de cas communs et amène alors à une sous-estimation de nombre de cas total. Le phénomène inverse est observé en cas de dépendance négative. Cette condition d'indépendance découle directement de l'homogénéité des captures. En effet, l'homogénéité implique que les individus identifiés et non identifiés dans une première source ont la même probabilité d'être identifiés dans l'autre source : l'identification par les deux sources est donc indépendante. Il est cependant d'usage d'énoncer ces conditions séparément (IWGDMF, 1995).

## 4.3 Estimateurs utilisés dans le cadre de la capture-recapture à deux sources

Plusieurs estimateurs existent pour calculer le nombre total de cas. Parmi ces estimateurs, ceux qui ont été le plus fréquemment rencontrés dans la littérature et qui sont décrits par la suite sont les estimateurs dits de Lincoln-Petersen, Chapman et Chao (Brittain et Böhning, 2009 ; El Allaki *et al.*, 2015).

Tableau 1 – Tableau de contingence du croisement de deux sources de données

Base	B			Total
	Présence	Oui	Non	
A	Oui	$n_{11}$	$n_{21}$	$N_A$
	Non	$n_{12}$	$n_{22} = ?$	
	Total	$N_B$		$N = ?$

Soient :

-  $n_{11}$  le nombre de cas présent à la fois dans la base de données A et la base de données B : il correspond donc au nombre de cas communs.

- $n_{12}$  et  $n_{21}$  le nombre de cas présents uniquement, respectivement, dans la base B et la base A.
- $N_A$  correspond au nombre total de cas dans la base A tandis que  $N_B$  est le nombre total de cas dans la base B.

Les deux inconnues qui sont estimées via la méthode de capture-recapture sont le nombre de cas qui ne sont présents dans aucune des deux bases, noté  $n_{22}$  (estimé implicitement), et le nombre total de cas, noté  $N$ .

L'approche la plus simple pour estimer ces deux valeurs est celle qui a recours à l'estimateur dit de Lincoln-Petersen. Cette approche historique ayant déjà été appliquée avant Lincoln et Petersen, l'appellation plus générale MRR pour *Mark Recapture Ratio* est parfois préférée (Goudie et Goudie, 2007 ; McNair *et al.*, 2018). Cette méthode ne s'applique que pour la capture-recapture à deux sources. Elle propose une estimation de  $N$ , notée  $\hat{N}$ , sur la base d'un modèle de type hypergéométrique tel que :

$$\hat{N} = \frac{N_A N_B}{n_{11}}$$

Le MRR correspond à l'estimateur du maximum de vraisemblance de cette distribution (Chapman, 1951 ; Wittes, 1972). Une manière de comprendre ce résultat est de considérer que, sous réserve d'indépendance des sources, l'Odds Ratio (OR) du tableau de contingence regroupant le statut détecté ou non pour les deux sources (Tab.1) est proche de 1. Ainsi, comme cela est décrit par Brittain et Böhning (2009), on a :

$$OR = \frac{n_{11}n_{22}}{n_{21}n_{12}}$$

Par conséquent, sous hypothèse d'indépendance des sources, on peut affirmer que :

$$OR \approx 1 \quad \text{donc} \quad \frac{n_{11}n_{22}}{n_{21}n_{12}} \approx 1 \quad \text{ainsi} \quad \hat{n}_{22} = \frac{n_{21}n_{12}}{n_{11}}$$

Par la suite, étant donné qu'on a :

$$\hat{N} = n_{11} + n_{21} + n_{12} + \hat{n}_{22}$$

on obtient en remplaçant  $n_{12}$  et  $n_{21}$  par, respectivement,  $N_B - n_{11}$  et  $N_A - n_{11}$  :

$$\hat{N} = n_{11} + (N_A - n_{11}) + (N_B - n_{11}) + \frac{(N_A - n_{11})(N_B - n_{11})}{n_{11}}$$

Par simplification, on obtient :

$$\hat{N} = \frac{N_A N_B}{n_{11}}$$

Par ailleurs, en considérant qu'il y a homogénéité de capture, on peut supposer que la proportion d'individus « marqués » dans la population est égale à la proportion d'individus « marqués » dans le second échantillon, comme résumé dans Chao *et al.* (2001). Ainsi, si on considère par exemple la base A comme première source, les individus de la population qui sont « marqués » sont ceux qui sont présents dans cette base de données. La proportion d'individus « marqués » dans la base A s'élève donc à  $\frac{N_A}{N}$ . Lors de la deuxième « capture », c'est-à-dire en considérant la base B, les individus « marqués » sont ceux présents dans la base B et qui ont déjà été détectés au préalable dans la base A. Cela correspond donc au nombre de cas communs aux deux bases. La proportion d'individus

marqués vaut donc  $\frac{n_{11}}{N_B}$ . Étant donné l'hypothèse d'homogénéité des captures, on peut donc considérer que :

$$\frac{N_A}{N} = \frac{n_{11}}{N_B}$$

Cela suppose également qu'il n'y a pas eu de mouvements de population entre la première et la deuxième « capture » pour considérer que la valeur de N est constante entre les deux captures et permettre l'écriture de cette égalité : on se situe donc dans une population dite close (McNair *et al.* 2018). Ainsi, nous aboutissons à la même conclusion que précédemment, à savoir :

$$\hat{N} = \frac{N_A N_B}{n_{11}}$$

Le MRR présente néanmoins un biais lorsque les effectifs sont petits et il ne peut pas être utilisé lorsque la probabilité d'obtenir des cas communs entre les deux sources est nulle. En effet, le MRR n'est pas défini lorsque  $n_{11}$  vaut 0. Un estimateur a été proposé par Chapman (1951) pour pallier cela en proposant une nouvelle expression de  $\hat{N}$  telle que :

$$\hat{N} = \frac{(N_A + 1)(N_B + 1)}{n_{11} + 1} - 1$$

Cela revient à considérer que (Brittain et Böhning, 2009) :

$$\hat{n}_{22} = \frac{n_{21}n_{12}}{n_{11} + 1}$$

L'estimateur de Chapman répond aux mêmes hypothèses que le MRR, notamment l'indépendance des sources et l'homogénéité de capture. Cet estimateur est celui qui est le plus communément utilisé dans le cadre de la capture-recapture à deux sources.

Dans le cas où il y aurait hétérogénéité de capture et dépendance entre les sources, les estimateurs présentés précédemment ne sont plus valables. Un estimateur adapté à ce cas de figure est proposé dans Chao (1987). Celui-ci permet le calcul d'une valeur minorant la valeur réelle de N. En notant  $f_1$  le nombre d'individus qui n'ont été détectés que par une seule source et S le nombre total d'individus détectés par les sources on a :

$$\hat{N} = S + \frac{f_1^2}{2n_{11}}$$

Cette écriture de l'estimateur de N est valable lorsque le nombre de sources est important (Chao, 1987 ; Farcomeni et Dotto, 2021). Ce cas de figure est souvent rencontré en écologie, contexte où cet estimateur a été développé en premier lieu, mais rarement en épidémiologie. Dans leur article de 2009, Brittain et Böhning formulent une expression de l'estimateur de Chao spécifiquement pour la capture-recapture à deux sources avec :

$$\hat{N} = S + \frac{f_1^2}{4n_{11}} \quad \text{Où } f_1 = N_A + N_B - 2n_{11} \text{ et } S = f_1 + n_{11}$$

Brittain et Böhning recommandent l'utilisation de l'estimateur de Chao après avoir comparé ses performances avec l'estimateur de Chapman sur des données réelles et des simulations. Leurs résultats montrent que cet estimateur présente un biais plus faible que l'estimateur de Chapman si les deux sources ne sont pas indépendantes. Cependant, ce n'est pas le cas lorsque les hypothèses

d'indépendance des sources et d'homogénéité de capture sont respectées, l'estimateur de Chapman donnant alors une estimation non biaisée de N. Par ailleurs, l'estimateur de Chao donne une estimation de N et non plus d'un minorant de N lorsqu'il y a homogénéité de capture (Chao et Colwell, 2017).

D'autres méthodes existent lorsque la capture-recapture est appliquée pour plus de deux sources. Il s'agit notamment de croiser les sources deux à deux ou de recourir à l'écriture de modèles log-linéaires et choisir le meilleur modèle selon des critères de sélection (critère d'information d'Akaike (AIC), critère d'information bayésien (BIC) etc.) (IWGDMF,1995). Ces modèles sont plus performants mais les résultats sont parfois plus difficiles à interpréter (Hook et Regal, 1995). Étant donné que nous ne disposons uniquement que de deux sources de données pour les cas graves de grippe admis en réanimation, la présente étude a eu recours à la méthode de capture-recapture à deux sources.

#### **4.4 Sources de données utilisées pour l'analyse de l'exhaustivité de la surveillance des cas sévères de grippe admis en réanimation par capture-recapture**

Pour notre analyse de l'exhaustivité de la surveillance des cas sévères de grippe par capture-recapture sur la saison 2021-2022, deux sources de données étaient disponibles. La première source est constituée par les données du système que l'on cherche à évaluer, c'est-à-dire les données de la surveillance des gripes sévères en réanimation (SGS) obtenues à partir des saisies des fiches cliniques transmises aux CR par les services de réanimation, tel que décrit précédemment (Section 1.3.1). La deuxième source est, elle, constituée par la base de données du Programme de Médicalisation des Systèmes d'Information (PMSI) en médecine, chirurgie, obstétrique et odontologie (MCO).

Le PMSI-MCO est un système d'information recueillant tous les actes médicaux en médecine, chirurgie, obstétrique et odontologie réalisés par les établissements de santé, aussi bien publics que privés. Les données concernant le patient, l'établissement de prise en charge et la nature des actes médicaux réalisés sont renseignés sur une plateforme sécurisée (e-PMSI). Le principal objectif de ce système est de répertorier les actes médicaux réalisés par un établissement de santé ainsi que la facturation associée dans le cadre de la tarification à l'activité (T2A). Cette dernière définit les financements alloués à chaque établissement de santé selon le nombre et la nature des actes médicaux réalisés (Ministère de la Santé). Chaque activité enregistrée donne lieu à la création d'un Résumé de Sortie Standardisé (RSS) composé d'un ou plusieurs Résumé d'Unité Médicale (RUM). Les RUM comportent toutes les informations concernant le séjour d'un patient dans une unité médicale. Les RSS anonymisés appelés Résumé de Sortie Anonyme (RSA) sont transmis aux Agences Régionales de Santé (ARS). Un code suivant les règles de la Classification Internationale des Maladies (CIM-10) est attribué à chaque diagnostic. Les séjours sont également classés selon des Groupes Homogènes de Malades (GHM). Le codage de chaque séjour est réalisé par le département d'information médicale (ATIH, 2022b).

L'utilisation de ces bases de données est définie par un cadre légal. Des autorisations doivent être accordées par la Commission Nationale de l'Informatique et des Libertés (CNIL) pour conduire la surveillance des gripes sévères permettant de constituer la base SGS et pour l'extraction des données de la base PMSI-MCO (Protocole Ecsir 2018, 2018).

## II – Travail personnel

### 1. Rappel du contexte, problématisation et objectifs

La surveillance des cas de grippe sévère admis en réanimation est une composante essentielle de la surveillance de la grippe en France. Les données qui en sont issues permettent entre autres l'évaluation de la sévérité des épidémies de grippe et la description des cas graves, essentielles pour mieux comprendre la maladie et appuyer les mesures de lutte et de prévention. Son évaluation est tout aussi centrale afin d'en connaître les performances, permettre de porter un regard critique sur les indicateurs issus des données de surveillance et améliorer le réseau de surveillance.

Cependant, après la réalisation d'une évaluation de l'exhaustivité de la surveillance des gripes sévères *via* les études Ecsir sur les saisons grippales 2009-2013, 2017-2018 et 2018-2019, aucune étude d'exhaustivité n'a été réalisée depuis la pandémie de COVID-19.

Ainsi, l'analyse que nous avons menée sur la saison 2021-2022 est la première depuis la pandémie et la première à être finalisée depuis le passage en surveillance sentinelle. Sa réalisation a été considérée comme prioritaire, notamment pour évaluer la qualité du réseau après trois ans de pandémie de COVID-19 qui ont eu un impact important sur l'organisation de la surveillance et l'offre et le recours aux soins. Cette étude a donc cherché à répondre à la question suivante : Quelle est l'exhaustivité de la surveillance des cas de grippe sévère admis en réanimation pour la saison grippale 2021-2022 ?

Pour cela, l'objectif principal de l'étude a consisté en l'estimation de l'exhaustivité de la surveillance des cas graves de grippe admis en réanimation pour la saison 2021-2022. Nous avons évalué l'exhaustivité à la fois de manière globale et en restreignant aux cas déclarés par les services sentinelles. Nos objectifs secondaires étaient de discuter des modifications pouvant être apportées au protocole des études Ecsir afin d'améliorer les estimations et apporter des résultats plus précis ainsi que de discuter de la validité des méthodes employées et la fiabilité des résultats.

### 2. Matériel et méthodes

#### 2.1 Accès et description des bases de données SGS et PMSI

Afin de mener notre étude d'exhaustivité de la surveillance des cas graves de grippe admis en réanimation, une première étape a consisté à avoir accès ou extraire les bases de données nécessaires à l'application de la méthode de capture-recapture à deux sources.

##### 2.1.1 Données de la surveillance des gripes sévères (SGS) et de la base des services de réanimation

La base de données de la surveillance des gripes sévères (SGS) est alimentée par les CR de Santé publique France à partir des fiches cliniques transmises par les services de réanimation participant au réseau (Annexe 1). La compilation des données recueillies par les CR permet l'obtention de la base nationale de cas de grippe sévère admis en réanimation pour la saison 2021-2022. Une base regroupant des services de réanimation ayant participé à la surveillance a également été mise à disposition dans l'objectif de la fusionner avec chacune des bases et d'attribuer à chaque cas des informations sur le service de réanimation correspondant.

Les informations présentes dans la base SGS sont l'identifiant du service qui a pris en charge le cas, des informations sur le médecin ayant rempli la fiche de signalement, la date de naissance du cas, son pays de naissance, son sexe, son département de domicile, sa commune, son arrondissement, des informations sur les antécédents de Covid-19, le statut vaccinal pour la grippe et la Covid-19 ainsi que sur des résultats de prélèvements, le motif d'admission, les éventuelles comorbidités et co-infections du cas, le type d'assistance ou aide ventilatoire, la date d'hospitalisation, la date d'admission en réanimation, la date de début des signes et la date de sortie/décès/transfert.

Concernant la liste des services de réanimation, les variables disponibles sont la région de l'hôpital où se situe le service de réanimation, l'identifiant du service, le type de service, le nom de l'hôpital, le numéro du département, le type d'hôpital, la participation ou non au réseau de surveillance sentinelle, le numéro FINESS géographique du service.

### 2.1.2 Données issues du Programme de Médicalisation des Systèmes d'Information (PMSI)

L'accès à la base PMSI a nécessité une demande d'autorisation d'extraction hors du Système National des Données de Santé (SNDS), réalisée par la Direction Appui, Traitements et Analyses de données de Santé publique France. Cette extraction a concerné les séjours hospitaliers avec un code diagnostique de grippe (code J09 à J11 selon la classification CIM-10) au sein d'une unité médicale (UM) de réanimation (UM 01, 02, 03), en diagnostic principal ou associé.

Cette extraction a concerné les périodes allant de la semaine 40 de 2021 à la semaine 20 de 2022 pour la France métropolitaine, la Guyane, la Réunion et Mayotte et de la semaine 40 de 2021 à de la semaine 30 de 2022 pour la Guadeloupe et la Martinique. Ces périodes ont été élargies par rapport aux périodes d'étude habituelles (semaine 40 ou 45 à la semaine 15 généralement) suite à une épidémie de grippe tardive. La période d'étude était encore plus étendue pour les départements des Antilles où l'épidémie y a débuté de manière particulièrement tardive.

Les variables dont nous disposons pour le PMSI sont les codes diagnostiques grippe, le type d'unité médicale (UM) de réanimation, la durée du séjour, le code identifiant unique, le mode de sortie, le code région d'hospitalisation, le code FINESS juridique de l'établissement, le code FINESS géographique de l'UM, le sexe du patient, l'âge du patient en années, le département de domicile du patient, la date d'hospitalisation, la date d'admission dans l'UM de réanimation et le code GHM du séjour.

L'accès à ces données SGS et PMSI ainsi que leur traitement dans le cadre de notre étude est permis par une autorisation permanente de la CNIL (délibération n°2017-305 du 7 décembre 2017/demande d'autorisation n°1929497) (CNIL, 2017).

## 2.2 Traitement des bases de données SGS et PMSI en amont de l'analyse

Une fois les données à disposition, une première étape a consisté au nettoyage de chaque base (SGS et PMSI) ainsi que leur fusion respective avec la base des services de réanimation.

### 2.2.1 Définition de la période d'étude

Dans la base SGS, seules les données portant sur des cas admis en réanimation pour les périodes sur lesquelles les données PMSI ont été extraites ont été retenues. Cependant, étant donné que l'extraction PMSI a été réalisée selon les dates d'hospitalisation et que nous avons décidé de travailler

avec des dates d'admission, la restriction des données sur cette période d'étude selon la date d'admission a également été réalisée dans le PMSI.

### 2.2.2 Ajout et retrait de variables

Tout d'abord, dans la base SGS, seules les variables en lien avec la grippe ont été retenues : les variables portant sur la COVID-19, du fait de la surveillance conjointe COVID-19/grippe, ont donc été supprimées pour alléger la base de données.

La variable âge n'étant pas disponible dans la base SGS, celle-ci a été calculée en réalisant la soustraction entre la date d'hospitalisation et la date de naissance et en convertissant le résultat en années. Le calcul de cette variable était nécessaire pour le croisement des sources. Le choix de la date d'hospitalisation comme référence de calcul pour l'âge se justifie par une volonté de cohérence avec la variable âge dans le PMSI qui est elle-même calculée à partir de la date d'hospitalisation (Loury *et al.*, 2017). Lorsque la date d'hospitalisation n'était pas disponible, la date d'admission en réanimation a été utilisée à sa place. Cette éventualité a été considérée étant donné qu'une description des caractéristiques de la variable « date d'hospitalisation » dans la base SGS nous a permis de détecter la présence de valeurs manquantes à hauteur de 2,9% après restriction des données sur la période d'étude définie. Cependant, ce nombre de valeurs manquantes restait assez faible et les valeurs prises par cette variable étaient cohérentes (antériorité par rapport à la date d'admission en réanimation vérifiée et absence de valeurs aberrantes), ce qui a donc permis son utilisation. Ces valeurs manquantes peuvent s'expliquer par le fait que l'ajout de la variable date d'hospitalisation dans la base SGS est récent (variable absente des précédentes études Ecsir), ce qui peut augmenter le risque d'erreurs ou d'oublis de saisie. La variable date d'admission ne possédait quant à elle aucune valeur manquante et les délais entre hospitalisation et admission en réanimation dans la base SGS étaient plutôt faibles (pour la valeur du délai entre les deux dates : médiane à 0 jours et moyenne à 1,8 jours). Elle s'est donc présentée comme une bonne variable de substitution.

Une variable « semaine d'admission » a également été créée, en utilisant cette fois-ci uniquement la date d'admission en réanimation. Les semaines ont été numérotées selon la norme ISO 8601.

Par la suite, la base SGS et la base des services de réanimation ont été fusionnées afin d'attribuer à chaque cas SGS les informations portant sur le service l'ayant pris en charge. Cela permet notamment d'introduire dans la base SGS la variable du FINESS géographique (présente initialement dans la liste des services), nécessaire pour le croisement des sources.

Une fois la fusion réalisée, les cas SGS ayant été reportés par des services de type « Unité de soins intensifs (USI) » et « Unité de soins continus (USC) » ont été supprimés. En effet, ces services sont différents des services de réanimation adulte et pédiatriques sur lesquels porte notre étude et ne correspondaient donc pas à nos critères d'inclusion.

Cette fusion avec la base des services permet également d'attribuer à chaque cas le nom de l'hôpital où celui-ci a été signalé. Une variable permettant de classer les cas selon leur hospitalisation en Centre Hospitalier Universitaire (CHU), Centre Hospitalier (CH) ou Autre a été créée.

Ces trois opérations ont également été réalisées dans la base PMSI à la différence que la sélection des services a été réalisée en gardant uniquement les cas où les variables geo01 ou geo13 (correspondant aux réanimation adultes et pédiatriques) étaient non vides.

Pour la base PMSI, une nouvelle variable date d'admission en réanimation a été calculée par la Direction Appui, Traitements et Analyses de données de Santé publique France en additionnant la date d'hospitalisation et la somme des durées passées dans les différentes unités médicales avant l'admission en réanimation. Le codage de la variable correspondant au département de domicile dans le PMSI pour les DROM a été harmonisé avec la notation dans la base SGS.

Une variable indiquant si un cas est issu d'un service sentinelle ou non a été créée dans les deux bases. Un cas a été considéré comme issu d'un service sentinelle si celui-ci a été déclaré par un service identifié comme sentinelle dans la base des services ou s'il provenait d'un service qui avait déclaré au moins un cas dans la base SGS mais non identifié comme sentinelle.

### 2.2.3 Gestion des valeurs manquantes

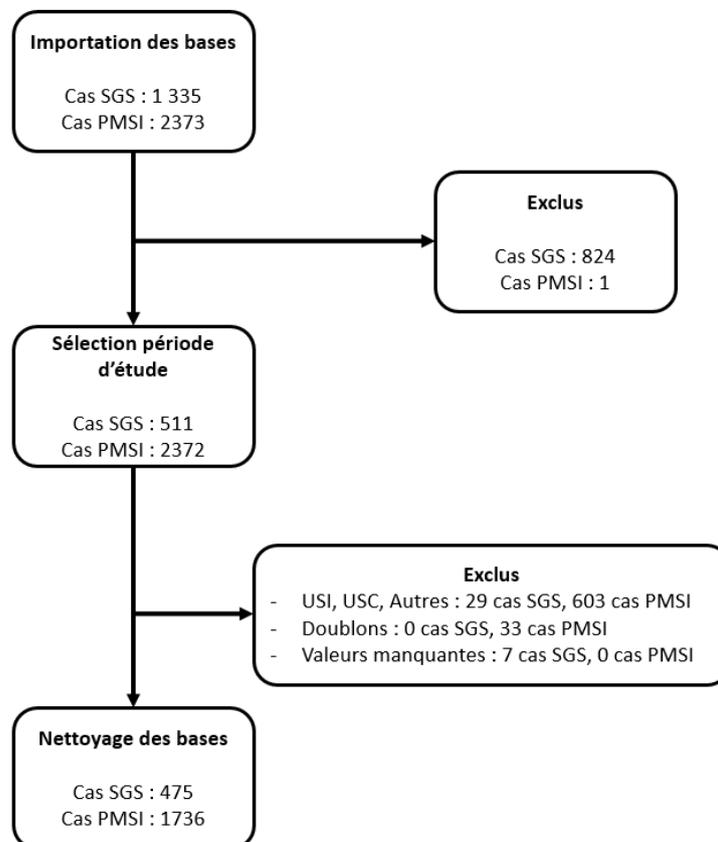
Les cas ayant des valeurs manquantes pour une des variables de croisement (voir II.2.3) (sauf pour le département de domicile) ont été retirés dans les deux bases.

### 2.2.4 Gestion des doublons

Au niveau de la base PMSI, il était possible d'avoir des doublons pour le numéro d'identification du séjour dans le cas où un patient avait été transféré vers un autre service ou avait été hospitalisé plusieurs fois au cours d'une même saison. Comme dans les précédentes études Ecsir, seul le séjour le plus ancien au sein de la saison a été gardé, à partir de la date d'hospitalisation comme référence. Cela permet de passer d'une base de séjours à une base de cas.

Le nombre de cas dans chaque source à chaque étape du nettoyage des bases de données est décrit dans la Figure 1. La base SGS utilisée au départ correspondait à une extraction couvrant la période d'août 2021 à février 2023. C'est la raison pour laquelle de nombreux cas SGS ont été exclus lors de l'étape de sélection de la période d'étude.

Figure 1 – Nombre de cas à chaque étape du nettoyage des bases de données SGS et PMSI préalable à l'analyse



Le traitement des bases de données ainsi que l'analyse des données ont été réalisés sous R 4.1.3.

## 2.3 Identification des cas communs par croisement des bases SGS et PMSI dans le cadre de la méthode de capture-recapture

Étant donné que nous ne disposions pas d'identifiant unique commun aux deux sources pour chaque cas, l'identification des cas communs entre les bases SGS et PMSI a été réalisée grâce à un croisement des deux bases sur cinq variables communes :

- Numéro FINESS géographique
- Sexe
- Âge
- Département de domicile
- Date d'admission en réanimation

Nous avons exécuté cinq requêtes différentes pour identifier les cas communs (Fig. 2).

### 2.3.1 Description des requêtes de croisement

La première requête (R1) a consisté à extraire les cas dans les deux bases présentant une égalité stricte pour chacune des variables de croisement

La deuxième requête (R2) a été rédigée comme R1 mais en enlevant l'égalité sur le département de domicile, cette variable pouvant s'avérer parfois moins fiable que les autres car généralement moins bien remplie.

La troisième requête (R3) a été rédigée comme R2 mais à la place d'avoir une égalité stricte pour l'âge, une différence de plus ou moins un an a été autorisée. Cette flexibilité sur l'âge avait été introduite dans les premières études Ecsir car la variable âge n'était pas calculée de la même manière dans la base SGS et PMSI (à partir de la date d'hospitalisation pour le PMSI et selon l'année d'admission pour certains cas de la base SGS) (Loury *et al.*, 2017). Dans notre cas, la variable âge n'était pas directement disponible dans la base SGS. Son calcul a donc été réalisé à partir de la date d'hospitalisation pour être en accord avec la manière dont cette variable a été calculée dans le PMSI. Cependant, le calcul a parfois été réalisé avec la date d'admission en réanimation lorsque la variable date d'hospitalisation présentait des valeurs manquantes. Il est donc possible qu'il puisse y avoir des différences d'âge pour un même patient entre les deux bases si celui-ci voit son âge augmenter d'un an entre ces deux dates. Nous avons donc laissé la flexibilité de plus ou moins un an pour l'âge pour prendre en compte cette éventualité.

La quatrième requête (R4) sélectionne les cas avec une égalité stricte pour le numéro FINESS géographique, le sexe et l'âge mais autorise une flexibilité pour la date d'admission en réanimation. Dans les précédentes études Ecsir, des différences allant de -1 jour à +7 jours entre les variables de date étaient acceptés. L'introduction de cette flexibilité s'expliquait par le fait que les variables de date qui étaient prises en compte pour la base SGS et la base PMSI n'étaient pas définies de la même manière. En effet, seule la date d'hospitalisation était disponible dans le PMSI et la base SGS ne comportait uniquement que la date d'admission en réanimation. Une analyse de sensibilité menée dans le cadre de l'étude Ecsir portant sur la période 2009-2013 a montré que lorsque l'on décrit la distribution du nombre de cas communs identifiés par la quatrième requête sur un intervalle allant de -10 jours à +10 jours, 96% de ces cas communs présentaient des différences de dates allant de -1 à +7 jours (Loury *et al.*, 2017).

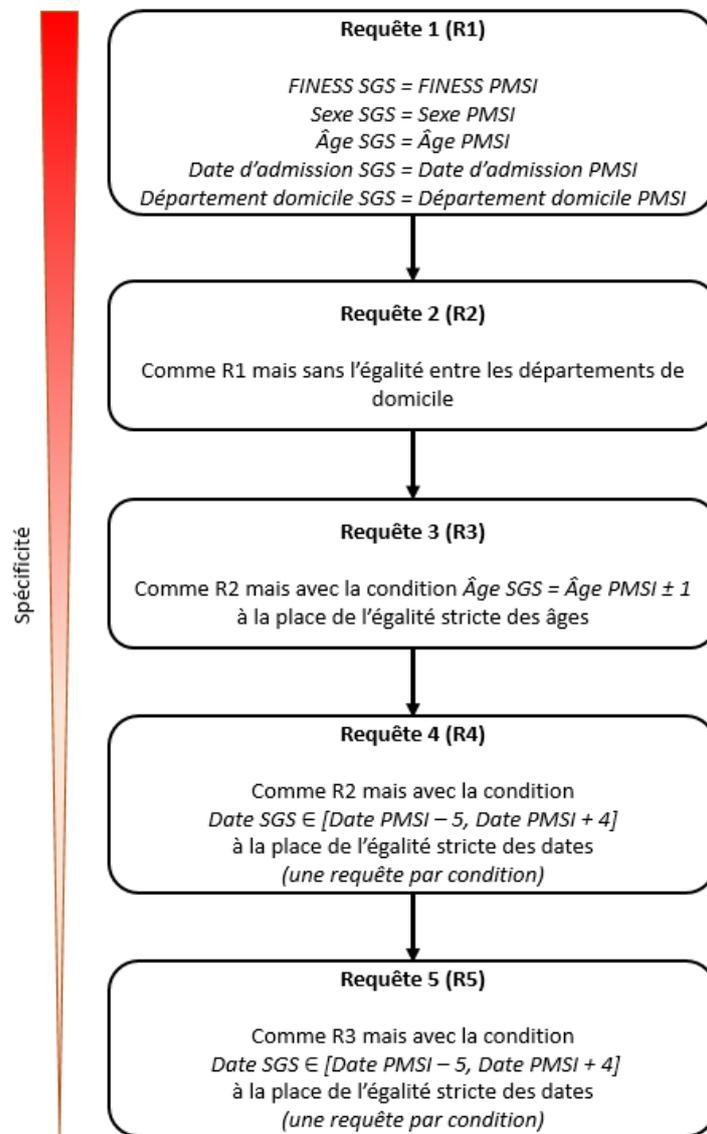
Étant donné que nous disposions de la date d'admission dans les deux bases, nous nous attendions à ce que les écarts de date sur les cas communs identifiés soient plus faibles et essentiellement liés à des erreurs de saisie. Nous avons réalisé à nouveau une analyse de sensibilité entre -10 jours et +10 jours pour identifier l'intervalle le plus adapté à notre étude. Étant donné que, dans notre cas, nous

avons la possibilité de tester des égalités entre des dates définies de la même manière (date d'admission pour les deux bases), autoriser une différence trop large entre les dates pouvait amener à prendre en compte des faux cas communs (personnes ayant le même âge et sexe, hospitalisées au même endroit mais ayant été admises en réanimation avec plusieurs jours de différence).

En autorisant une différence de dates d'admission en réanimation allant de -10 jours à +10 jours, nous avons pu identifier 44 cas communs pour la requête R4. Plus de 95% de nos cas présentaient des différences de date allant de -5 jours à +4 jours. Nous avons alors autorisé des différences de dates comprises dans cet intervalle dans les quatrièmes et cinquièmes requêtes.

La cinquième et dernière requête (R5) reprend la requête R4 mais autorise à nouveau une différence de plus ou moins un an pour l'âge.

Figure 2 – Description et ordre d'exécution des requêtes de croisement



L'ordre d'exécution des requêtes allait de R1 à R5, c'est-à-dire du plus spécifique au moins spécifique (Fig. 2). Pour les requêtes R4 et R5, une requête différente a été écrite pour chaque différence de date autorisée (une pour +1 jour, une pour -1 jour et ainsi de suite, dans cet ordre). À chaque fois qu'une flexibilité était accordée sur une variable, l'égalité stricte n'était pas testée au sein

de la requête en question car ces cas avaient déjà été identifiés par les requêtes précédentes (R1 et R2 notamment).

De plus, à la fin de chaque requête, une vérification de la présence de doublons dans les cas communs identifiés a été réalisée à partir des identifiants uniques propres à chacune des deux bases. Il est en effet possible qu'un cas dans une base puisse croiser avec plusieurs cas de l'autre base, notamment lorsqu'une flexibilité est accordée sur certaines valeurs. Dans ce cas, un couple SGS-PMSI était choisi aléatoirement et désigné comme cas commun. Les cas identifiés par une requête n'étaient plus pris en compte dans les requêtes suivantes. Seuls les cas n'ayant pas croisé et les cas ayant été détectés comme doublons et non désignés comme communs pouvaient à nouveau être considérés dans la requête suivante.

### 2.3.2 Comparaison avec les requêtes de croisement des précédentes études Ecsir

Notre méthodologie pour le croisement des sources reprenait exactement celle élaborée pour les études Ecsir 2017-2018 et 2018-2019 à la seule différence de l'utilisation de la date d'admission en réanimation comme variable de croisement. Disposant des deux types de dates pour chacune des deux bases, notre choix s'est tourné vers l'utilisation de la date d'admission pour écrire les requêtes de croisement. Cela se justifie par le fait que cette variable est considérée comme plus fiable (moins d'erreurs de saisie et pas de valeurs manquantes dans la base SGS). De plus, lorsque les requêtes étaient exécutées en prenant en compte uniquement la date d'hospitalisation ou en prenant en compte les deux variables de date (croisement avec 6 variables communes au lieu de 5), il y avait moins de cas communs identifiés que lorsque seule la date d'admission était considérée. En outre, davantage de cas communs étaient récupérés dès la première requête en prenant en compte uniquement la date d'admission, ce qui témoignait d'une meilleure spécificité du croisement.

L'ordre d'écriture des requêtes R4 et R5 était également différent. Dans les précédentes études, l'ordre des requêtes avec une flexibilité pour la date était +1 jour, +2 jours, ..., +7 jours puis -1 jour car la date d'admission SGS avait plus de chances d'être postérieure à la date d'hospitalisation PMSI. Cela permettait notamment de limiter les erreurs de croisement. Étant donné que nous avons croisé nos bases sur la date d'admission pour les deux bases, nous avons plutôt opté pour l'ordre +1 jour, -1 jour, +2 jours, -2 jours etc.

Chaque requête a été rédigée sous R en langage SQL via le package sqldf (Grothendieck, 2017). Nous n'avons donc pas eu recours à l'algorithme de croisement utilisé jusque-là dans les précédentes études Ecsir écrit sous Stata.

### 2.3.3 Obtention de la base de données finale

Par la suite, une variable comportant les résultats des requêtes de croisement a été créée pour que chacun des cas puisse être caractérisé selon son statut de croisement. Cette variable prend la valeur « commun » pour les cas ayant croisé, « sgs » pour ceux étant présents uniquement dans la base SGS et « pmsi » pour ceux présents uniquement dans le PMSI. Les valeurs des variables de croisement gardées pour les cas communs dans la base finale étaient les valeurs du PMSI, celles-ci étant considérées comme plus fiables que celles de la base SGS.

La base de données finale comprenait les cas communs et les cas qui ont été identifiés uniquement dans la base SGS ou dans le PMSI.

## 2.4 Calcul de l'exhaustivité par la méthode de capture-recapture

Une fois les conditions d'application vérifiées et la base de données finale obtenue, une estimation du nombre de cas total  $N$  a été réalisée en appliquant la méthode de capture-recapture.

### 2.4.1 Vérification des conditions d'application de la méthode de capture-recapture

Les conditions d'application de la méthode de capture-recapture étaient pour la plupart respectées.

L'identification des cas par la surveillance SGS et le PMSI se fait en France (métropole et DROM) et seules les données SGS correspondant aux périodes sur lesquelles ont été extraites les données PMSI ont été considérées. Ainsi les cas étaient détectés selon les mêmes conditions spatio-temporelles dans les deux sources.

La population que l'on cherche à estimer pouvait être considérée comme étant globalement close. L'identification des cas se repose sur le diagnostic de grippe : il suffit que le diagnostic ait été posé à un moment  $t$  inclus dans notre période d'étude pour que le cas soit identifiable par les deux sources. Les naissances, les décès et les déplacements hors de la zone géographique d'étude n'affectent donc pas l'identification de cas.

Le respect de la condition énonçant que tout cas identifié doit être un vrai cas sous-entend généralement que les deux sources ont la même définition de cas. D'après Loury *et al.* (2017), cela peut se traduire par le fait que la proportion de cas confirmés biologiquement est similaire dans les deux sources. Cependant, alors que les précédentes définitions de cas comportaient la possibilité d'inclure les formes cliniques graves dont les symptômes et l'anamnèse évoquent un diagnostic de grippe (cas probable) (Pivette *et al.*, 2019), la nouvelle définition de cas utilisée par la surveillance SGS est exclusive aux cas confirmés biologiquement (Annexe 1). D'après la classification CIM-10, les codes allant de J09 à J11 concernent les gripes avec ou sans diagnostic biologique (ATIH, 2022a). Néanmoins, ce codage n'était pas disponible dans notre base PMSI, ce qui a rendu impossible le calcul de la proportion de cas non testés virologiquement et de voir si celui-ci était négligeable. Cependant, d'après Loury *et al.* (2017), le codage grippe PMSI concerne principalement les cas confirmés biologiquement pour limiter le nombre de cas non confirmés dans la base hospitalière. Nous avons alors décidé de supposer que cette condition était respectée.

Les requêtes de croisement ont été écrites de manière à être les plus spécifiques possibles pour identifier les vrais cas communs tout en autorisant une certaine flexibilité sur les dernières requêtes pour maximiser le nombre de cas et permettre d'identifier tous les cas communs. Nous avons donc considéré que la condition d'identification de tous les vrais cas communs était vérifiée.

Les hypothèses d'homogénéité de capture et d'indépendance des sources sont souvent posées dans le cadre des études de capture-recapture. L'homogénéité implique l'indépendance mais il est d'usage d'énoncer les deux hypothèses séparément (IWGGMF, 1995 ; Chao *et al.*, 2001). Cependant, ces deux conditions sont difficiles à vérifier quantitativement pour la capture-recapture à deux sources. Dans le contexte d'études épidémiologiques, il est même peu probable que ces deux conditions soient respectées. Dans le cadre de notre étude, nous pouvons par exemple considérer que la probabilité d'être détecté par la surveillance SGS dépend fortement de si le cas est admis dans un service de réanimation sentinelle. Des facteurs sociodémographiques ou géographiques peuvent également avoir un impact sur les probabilités de capture. Toutefois, l'hétérogénéité peut être appréhendée par la stratification des analyses selon les variables qu'on estime comme pouvant être source d'hétérogénéité pour réduire les biais (IWGDMF, 1995 ; Papoz *et al.*, 1996 ; Tilling, 2001).

Étant donné qu'il n'est pas possible de se prononcer de manière certaine sur le respect de ces deux conditions, nous avons comparé plusieurs estimateurs plus ou moins sensibles aux hypothèses d'homogénéité et d'indépendance pour pouvoir ensuite discuter qualitativement de la validité des hypothèses.

### 2.4.2 Choix de l'estimateur du nombre de cas total

Lorsque les hypothèses d'indépendance des sources et d'homogénéité de capture ne peuvent plus être posées, les estimations du nombre de cas total  $N$  données par le MRR et l'estimateur de

Chapman sont biaisées (Gallay *et al.*, 2002). L'estimateur de Chao est donc plus adapté à cette situation car moins sensible à ces hypothèses (Brittain et Böhning, 2009).

Dans le cadre de notre étude, les estimateurs de Chapman et Chao ont tous deux été utilisés, l'objectif étant de comparer les résultats des estimations, comme cela a par exemple déjà été réalisé par La Ruche *et al.* (2013). L'estimateur de Chapman a été privilégié par rapport au MRR car une première description de nos données a mis en évidence la présence d'effectifs assez faibles dans certaines strates. Les résultats d'exhaustivité ont également été présentés sous forme d'un intervalle dit de plausibilité, noté ici  $I_{Ch-BCS}$ , reprenant la méthodologie utilisée par Hassel *et al.* (2011), avec la borne inférieure de l'IC 95% de l'estimateur de Chao comme borne inférieure et une borne supérieure définie à partir d'une situation idéale (« *best case scenario* », noté BCS ici) où les tous les cas peuvent être retrouvés dans au moins une des deux sources. Les bornes sont inversées lorsque les valeurs présentées sont celles du nombre de cas total estimé (et c'est la borne supérieure de l'IC95% de l'estimation de Chao qui est considérée dans ce cas). Dans le scénario idéal, on considère que  $N = N_{SGS} + N_{PMSI} - n_{11}$ , c'est-à-dire le nombre de cas total identifié par les deux sources. On peut alors généraliser avec  $N \geq N_{SGS} + N_{PMSI} - n_{11}$ . Ce minorant de N devient un majorant lorsqu'on parle de l'exhaustivité.

Le MRR étant l'estimateur utilisé dans les précédentes études Ecsir, celui-ci a également été utilisé dans le but de pouvoir comparer les résultats donnés par celui-ci avec les autres estimations, notamment en présence de faibles effectifs. Les valeurs de N sont arrondies à l'entier inférieur.

### 2.3.3 Méthode de calcul de l'exhaustivité et des intervalles de confiance associés

Une fois le nombre de cas total N estimé (noté  $\hat{N}$  ici), le calcul de l'exhaustivité de chacune des bases (en pourcent) a été réalisé selon l'expression suivante :

$$Exhaustivité_{base} = \frac{N_{base}}{\hat{N}} * 100$$

Cette estimation de l'exhaustivité s'est accompagnée du calcul des intervalles de confiance à 95% (IC 95%) en ayant recours à la méthode du bootstrap paramétrique. Le choix d'opter pour une méthode différente du calcul des IC par l'approximation normale ( $IC(N) = \hat{N} \pm 1,96\sqrt{Var(\hat{N})}$ ) se justifie par le fait que d'après le travail de revue de l'International Working Group for Disease Monitoring and Forecasting (IWGDMF) (1995), la distribution de  $\hat{N}$  dans la littérature présente généralement un coefficient d'asymétrie (*skewness*) non nul, contrairement à une distribution suivant une loi normale. Le recours à l'approximation normale n'est donc pas le plus appropriée pour le calcul des IC. L'approche bootstrap constitue une des alternatives mentionnées par l'IWGDMF.

La méthode du bootstrap paramétrique consiste à réaliser un nombre r de ré-échantillonnages dans une distribution (tirage avec remise). La distribution que nous avons choisie est une distribution binomiale  $B(n, p)$  de paramètre  $n = n_{commun} = n_{11}$  et  $p = \frac{n_{11}}{N_{PMSI}}$ . Cela permet de modéliser une distribution du nombre de cas communs identifiés. Nous avons considéré que la source dont on veut évaluer l'exhaustivité (SGS) correspond à la « capture » initiale et que la source avec laquelle on la croise (PMSI) correspond à la « recapture », même s'il n'y a en général pas de temporalité dans la capture-recapture appliquée aux populations humaines (Chao *et al.*, 2001). Un cas ne peut être désigné comme commun qu'après l'étape de « recapture ». C'est la raison pour laquelle le paramètre p est défini comme valant  $p = \frac{n_{11}}{N_{PMSI}}$  avec  $N_{PMSI}$  au dénominateur. Les fonctions permettant de calculer les IC pour un estimateur donné ont été écrites en s'inspirant de la structure de la fonction ciChapman du package *recapr* (Tyers, 2021), qui utilise également une loi binomiale pour modéliser la distribution du nombre de cas communs. Une condition a été ajoutée dans la fonction pour que le nombre de cas commun ne puisse pas être supérieur à l'effectif de la source avec le plus faible nombre de cas ( $n_{11} \leq \min(N_{SGS}, N_{PMSI})$ ).

Le nombre de ré-échantillonnages  $r$  a été fixé à 10 000 (valeur par défaut dans la fonction `ciChapman`). Des valeurs de  $r$  inférieures à 10 000 ont été testées dans un but d'optimisation avec des valeurs de  $r$  allant de 1000 à 9000 avec un pas de 1000. L'objectif était de savoir s'il était possible d'obtenir les mêmes résultats pour moins de ré-échantillonnages. Aucune différence de temps de calcul n'a été perçue qualitativement. Quelques différences ont été observées sur la valeur des estimations des IC par rapport aux estimations obtenues avec  $r$  valant 10 000. La valeur du paramètre  $r$  est donc restée fixée à 10 000.

Par la suite, une distribution du nombre total de cas  $\hat{N}$  a pu être obtenue en calculant  $\hat{N}$  par l'estimateur considéré pour chacune des valeurs de  $n_{11}$  échantillonnées (10 000 valeurs en tout). Enfin, une distribution des valeurs d'exhaustivité pour chacune des bases a été obtenue en la calculant à partir des valeurs de  $\hat{N}$  calculées précédemment. Les quantiles à 2,5% et 97,5% de ces distributions de la valeur de l'exhaustivité permettent d'obtenir respectivement la borne inférieure et la borne supérieure de l'intervalle de confiance. Une même graine a été définie avant l'exécution de chaque ligne de code impliquant un calcul d'IC bootstrap pour assurer la reproductibilité des résultats.

L'avantage du recours à la méthode bootstrap est que celle-ci est moins restrictive au niveau des hypothèses nécessaires à son application par rapport à un calcul des IC avec l'approximation normale. Les intervalles de confiance résultant de la méthode bootstrap sont généralement plus larges que ceux issus de l'approximation normale. Cela diminue les chances d'observer des différences significatives mais lorsque de telles différences sont observées, il y a de plus grandes chances pour que cette significativité soit avérée.

## 2.5 Analyses réalisées

Une fois le nombre total de cas de grippe sévères admis en réanimation estimé, une analyse de l'exhaustivité à l'échelle nationale et métropole a été menée.

Par la suite, une analyse stratifiée selon les classes d'âge et les régions a été réalisée. Les classes d'âge considérées étaient : 0-1 an, 2-17 ans, 18-64 ans et 65 ans et plus. Ce découpage inhabituel dans le cadre de la surveillance de la grippe avait pour objectif d'inclure la classe d'âge des 2-17 ans pour pouvoir disposer de données spécifiques à cette classe d'âge suite à l'ouverture de la vaccination contre la grippe pour les enfants de 2 à 17 ans révolus sans comorbidités par la Haute Autorité de Santé (HAS) en février 2023. Concernant l'analyse stratifiée par région, la présence d'effectifs trop faibles dans certaines de ces régions nous a amené à renoncer à estimer l'exhaustivité pour ces régions en question et de simplement présenter le nombre de cas. Cette décision a été prise en accord avec les CR concernées. Les régions où l'exhaustivité n'a pas été estimée étaient : la Guadeloupe, la Martinique, la Guyane et la Corse. Les données de ces régions ont été gardées dans les autres types d'analyses. Des analyses complémentaires en stratifiant par type d'hôpital et par période de temps ont également été menées. Chaque analyse a aussi été réalisée de manière globale et en restreignant aux services sentinelles.

Un calcul du taux d'incidence des cas de grippe sévère admis en réanimation pour cent mille habitants a également été réalisé en divisant le nombre de cas total estimé par les données de population de 2022 (INSEE, 2023).

Les sorties sous formes de graphiques ont été réalisées principalement à l'aide des packages `R ggplot2` (Wickham *et al.*, 2016) pour les sorties statiques et `plotly` (Sievert *et al.*, 2020) pour les sorties interactives incluses dans un tableau de bord sous format HTML qui sera mis à disposition des CR avec davantage de données spécifiques aux régions. Ce tableau de bord a été créé à l'aide du package `flexdashboard` (Sievert *et al.*, 2023). Les tableaux de données ont été réalisés en utilisant les packages `gtsummary` (Sjoberg *et al.*, 2021) et `flextable` (Gohel et Skintzos, 2022).

Les sorties sous forme de cartes ont, elles, été réalisées à l'aide du package `mapsf` (Giraud *et al.*, 2022). La méthode de découpage des catégories utilisée pour les cartes est celle de Jenks qui permet

de minimiser la variance intra-classe tout en maximisant la variance interclasse et ainsi créer des catégories homogènes et distinctes entre elles.

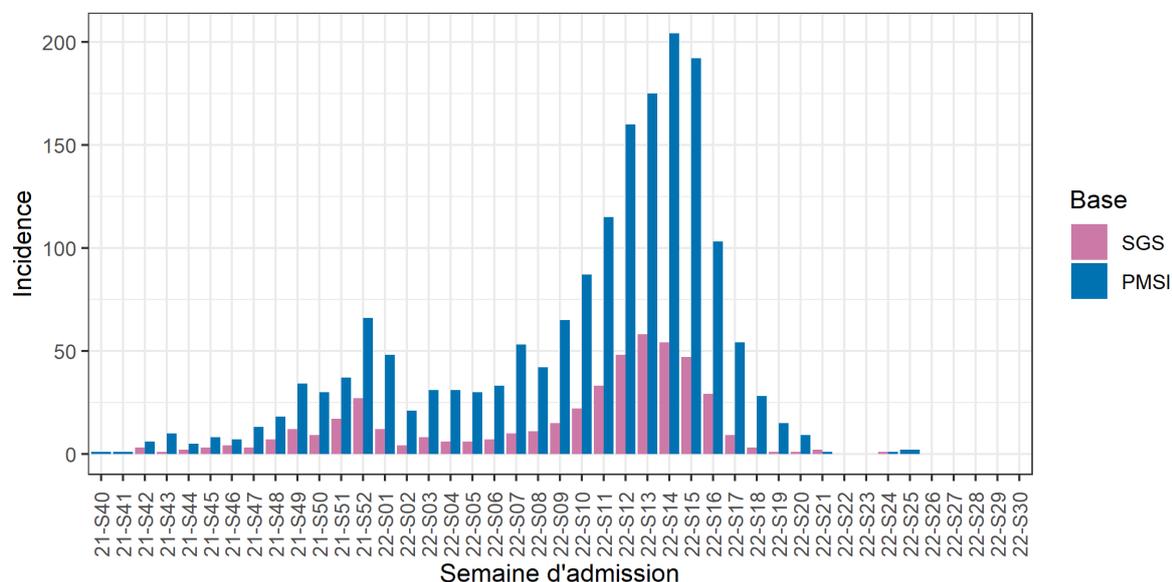
### 3. Résultats

#### 3.1 Description des bases SGS et PMSI

##### 3.1.1 Incidence des cas graves de grippe admis en réanimation par semaine

Tout d’abord, nous avons cherché à décrire la distribution temporelle de l’incidence des cas graves de grippe admis en réanimation pour chaque semaine et pour chacune des deux bases de données (Fig.3).

Figure 3 – Incidence des cas graves de grippe admis en réanimation par semaine d’admission en réanimation, France entière, Semaine 40 de 2021 à Semaine 30 de 2022. Les semaines sont numérotées dans le format Année-Semaine.



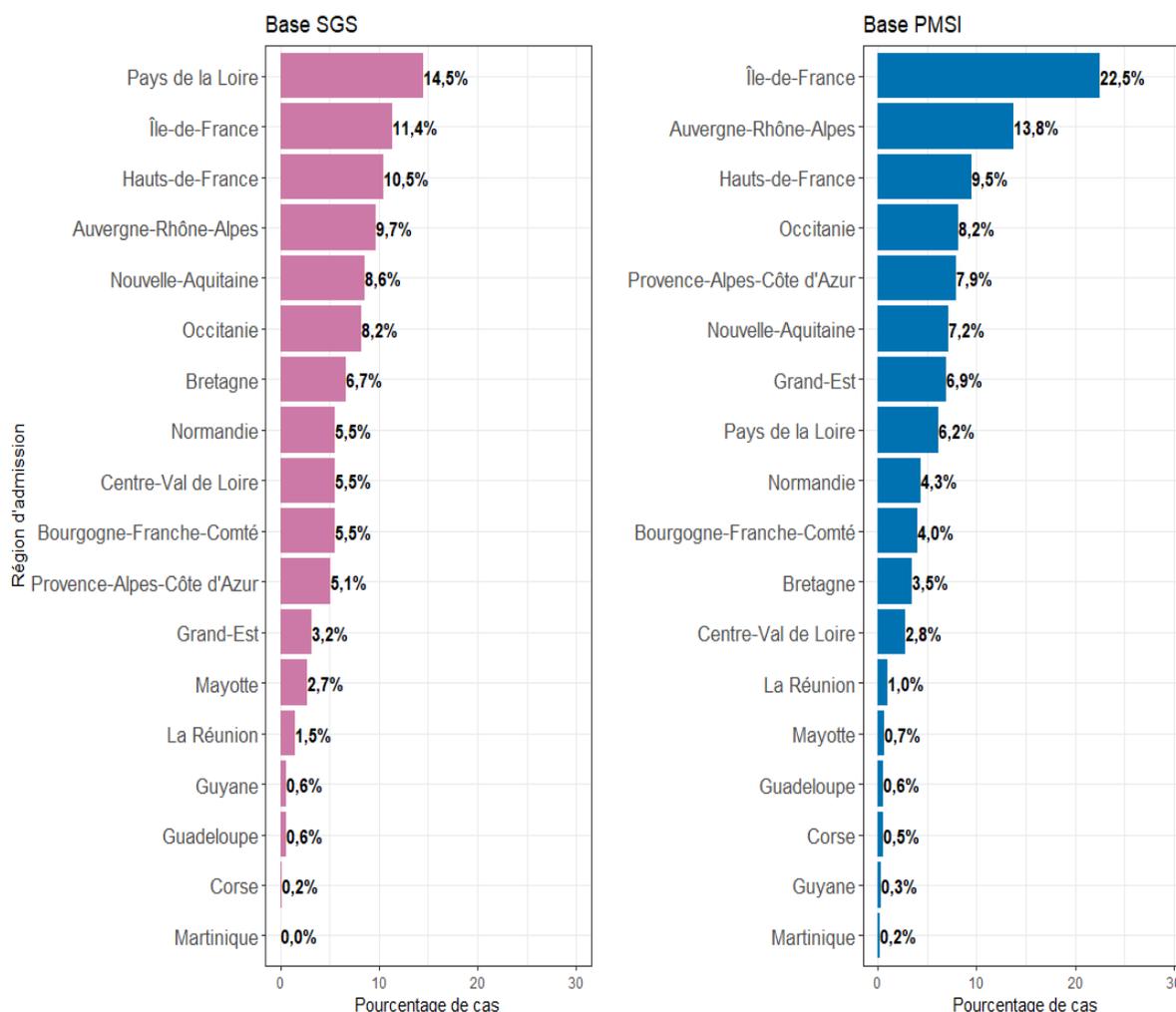
On observe que les courbes épidémiques obtenues à partir des données de chacune des sources présentait une allure similaire avec tout d’abord une augmentation du nombre de cas de la semaine 40 jusqu’à la semaine 52 de 2021, suivie d’une diminution puis d’une montée de l’incidence plus importante à partir de la semaine 09 de 2022. Le pic épidémique est atteint en semaine 13 de 2022 pour les données SGS et en semaine 14 pour les données PMSI. L’incidence diminuait ensuite progressivement jusqu’à la fin de l’épidémie en semaine 20. Les cas apparaissant au-delà de la semaine 20 correspondaient aux cas signalés en Guadeloupe et en Martinique où l’épidémie y a été plus tardive.

Au total, la base SGS comportait 475 cas et la base PMSI 1 736 cas.

##### 3.1.2 Répartition des cas par région d’admission selon les bases de données

A présent, lorsque l’on s’intéresse à la répartition des cas selon les régions, on remarque des proportions de cas signalés assez disparates selon les régions considérées. On observe également une répartition des cas par région différente entre les bases SGS et PMSI, même si les régions figurant dans celles ayant déclaré le plus et le moins de cas restaient globalement similaires dans les deux sources (Fig.4).

Figure 4 – Répartition des cas par région d'admission classé par ordre décroissant, par base de données, France entière, Semaine 40 de 2021 à Semaine 30 de 2022



Ainsi, pour la base SGS, la majorité des cas ont été signalés dans la région Pays de la Loire avec 14,5% des cas. Les DROM et la région Grand-Est représentaient un faible nombre de cas, en particulier la Martinique où aucun cas n'a été retenu après nettoyage de la base SGS.

Pour la base PMSI, c'est la région Ile-de-France qui concentrait le plus de cas au niveau national avec 22,5% des cas. Les pourcentages les plus faibles sont à nouveau observés au niveau des DROM où les pourcentages de cas étaient inférieurs ou égaux à 1%.

### 3.1.3 Caractéristiques sociodémographiques des cas graves de grippe admis en réanimation par source

Les deux bases de données présentaient certaines variables en commun comme l'âge, le sexe, la région d'hospitalisation ou encore l'unité médicale ayant pris en charge le patient. La plupart de ces variables intervenaient notamment pour permettre de déterminer les cas communs entre les deux bases. Certaines ont également été utilisées pour réaliser des analyses stratifiées. Ces variables ont été décrites pour chacune des bases dans le Tableau 2.

Tableau 2 – Caractéristiques des cas graves de grippe admis en réanimation pour les variables sexe, âge, classe d'âge, région, type d'hôpital et unité médicale pour chaque source, France entière, Semaine 40 de 2021 à Semaine 30 de 2022

<b>Variables</b>	<b>SGS<sup>1</sup></b>	<b>PMSI<sup>1</sup></b>	<b>p-valeur</b>
<b>Effectif total</b>	475	1736	
<b>Sexe</b>			0,7 <sup>2</sup>
Femme	215 (45,3%)	768 (44,2%)	
Homme	260 (54,7%)	968 (55,8%)	
<b>Âge (médiane, EI)</b>	57 (20,5 – 67)	60 (28,0 – 70)	0,009 <sup>3</sup>
<b>Classe d'âge</b>			0,062 <sup>2</sup>
0-1 an	41 (8,6%)	162 (9,3%)	
2-17 ans	67 (14,1%)	211 (12,2%)	
18-64 ans	206 (43,4%)	667 (38,4%)	
65 ans et plus	161 (33,9%)	696 (40,1%)	
<b>Région</b>			<0,001 <sup>4</sup>
Auvergne-Rhône-Alpes	46 (9,7%)	240 (13,8%)	
Bourgogne-Franche-Comté	26 (5,5%)	69 (4,0%)	
Bretagne	32 (6,7%)	60 (3,5%)	
Centre-Val de Loire	26 (5,5%)	48 (2,8%)	
Corse	1 (0,2%)	9 (0,5%)	
Grand-Est	15 (3,2%)	119 (6,8%)	
Guadeloupe	3 (0,6%)	10 (0,6%)	
Guyane	3 (0,6%)	5 (0,3%)	
Hauts-de-France	50 (10,5%)	165 (9,5%)	
Île-de-France	54 (11,4%)	390 (22,5%)	
La Réunion	7 (1,5%)	18 (1,0%)	
Martinique	0 (0,0%)	4 (0,2%)	
Mayotte	13 (2,7%)	12 (0,7%)	
Normandie	26 (5,5%)	75 (4,3%)	
Nouvelle-Aquitaine	41 (8,6%)	125 (7,2%)	
Occitanie	39 (8,2%)	142 (8,2%)	
Pays de la Loire	69 (14,5%)	107 (6,2%)	
Provence-Alpes-Côte d'Azur	24 (5,1%)	138 (7,9%)	
<b>Type d'hôpital</b>			<0,001 <sup>2</sup>
CHU	276 (58,1%)	890 (51,3%)	
CH	194 (40,8%)	698 (40,2%)	
Autre	5 (1,1%)	148 (8,5%)	
<b>Unité médicale</b>			>0,9 <sup>2</sup>
Réanimation adulte	377 (79,4%)	1376 (79,3%)	
Réanimation pédiatrique	98 (20,6%)	360 (20,7%)	

<sup>1</sup>n (%); Médiane (EI)

<sup>2</sup>test du khi-deux d'indépendance

<sup>3</sup>test de Wilcoxon-Mann-Whitney

<sup>4</sup> Test exact de Fisher avec p-valeur simulée (2000 répétitions)

On observe des distributions d'effectifs identiques ou quasi identiques pour les valeurs des variables « unité médicale » et « sexe » entre les deux bases de données. Une légère différence est observée au niveau de la variable âge avec les cas de la base SGS qui apparaissent un peu plus jeunes que ceux dans le PMSI. Cela s'est confirmé avec l'observation de différences de distribution des

effectifs dans les classes d'âge selon les bases, la classe majoritaire dans la base SGS étant les 18-64 ans alors que dans le PMSI ce sont les 65 ans et plus.

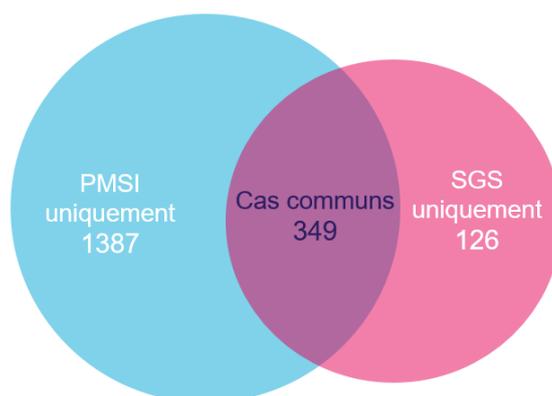
Le tableau 2 montre également que la répartition des effectifs selon les régions était plus homogène dans la base SGS que le PMSI. Concernant la variable type d'hôpital, on remarque que les cas issus des CHU étaient majoritaires dans les deux bases et qu'il y avait une proportion plus importante de cas déclarés par des hôpitaux classés dans Autre (cliniques privées, hôpitaux d'instruction militaire) dans le PMSI par rapport à la base SGS. Si l'on considère un seuil à 0,05, les variables région ( $p < 0,001$ ) et type d'hôpital ( $p < 0,001$ ) étaient significativement différentes entre les deux bases. Ceci était également le cas pour les données restreintes aux services sentinelles (Annexe 2).

### 3.2 Croisement des bases SGS et PMSI

#### 3.2.1 Résultats du croisement

Le croisement de la base SGS avec la base PMSI a été réalisé sur 5 variables communes (voir Section II.2.3). Les cas des deux bases remplissant les critères définis par nos requêtes à partir de ces 5 variables ont été considérés comme cas communs. Au total, 349 cas communs ont été identifiés par nos requêtes. Le nombre de cas présents uniquement dans la base SGS et la base PMSI étaient, respectivement, 126 cas et 1387 cas (Fig.5).

Figure 5 – Diagramme de Venn avec le nombre de cas communs aux deux bases et le nombre de cas spécifique à chaque base



Un tableau récapitulatif le nombre de cas récupérés par chaque requête est mis en annexe (Annexe 3).

#### 3.2.2 Caractéristiques sociodémographiques des cas graves de grippe admis en réanimation selon les résultats du croisement

Une description des valeurs de certaines variables communes en fonction du résultat de croisement pour les analyses globales est présentée dans le tableau suivant (Tab.3).

Tableau 3 - Caractéristiques des cas graves de grippe admis en réanimation pour les variables sexe, âge, classe d'âge, région, type d'hôpital et unité médicale selon les résultats du croisement, France entière, Semaine 40 de 2021 à Semaine 30 de 2022

Variabiles	Commun <sup>1</sup>	PMSI uniquement <sup>1</sup>	SGS uniquement <sup>1</sup>
<b>Effectif total</b>	349	1387	126
<b>Sexe</b>			
Femme	158 (45,3%)	610 (44,0%)	57 (45,2%)
Homme	191 (54,7%)	777 (56,0%)	69 (54,8%)

<b>Variables</b>	<b>Commun<sup>1</sup></b>	<b>PMSI uniquement<sup>1</sup></b>	<b>SGS uniquement<sup>1</sup></b>
<b>Âge (médiane)</b>	60 (26,0 – 69)	60 (29,0 – 71)	42 (15,5 – 62)
<b>Classe d'âge</b>			
0-1 an	29 (8,3%)	133 (9,6%)	12 (9,5%)
2-17 ans	46 (13,2%)	165 (11,9%)	21 (16,7%)
18-64 ans	140 (40,1%)	527 (38,0%)	66 (52,4%)
65 ans et plus	134 (38,4%)	562 (40,5%)	27 (21,4%)
<b>Région</b>			
Auvergne-Rhône-Alpes	39 (11,2%)	201 (14,5%)	7 (5,6%)
Bourgogne-Franche-Comté	22 (6,3%)	47 (3,4%)	4 (3,2%)
Bretagne	25 (7,2%)	35 (2,5%)	7 (5,6%)
Centre-Val de Loire	16 (4,6%)	32 (2,3%)	10 (7,9%)
Corse	1 (0,3%)	8 (0,6%)	0 (0,0%)
Grand-Est	13 (3,7%)	106 (7,6%)	2 (1,6%)
Guadeloupe	1 (0,3%)	9 (0,6%)	2 (1,6%)
Guyane	3 (0,9%)	2 (0,1%)	0 (0,0%)
Hauts-de-France	41 (11,7%)	124 (9,0%)	9 (7,1%)
Île-de-France	37 (10,6%)	353 (25,5%)	17 (13,5%)
La Réunion	4 (1,1%)	14 (1,0%)	3 (2,3%)
Martinique	0 (0,0%)	4 (0,3%)	0 (0,0%)
Mayotte	2 (0,6%)	10 (0,7%)	11 (8,7%)
Normandie	18 (5,2%)	57 (4,1%)	8 (6,3%)
Nouvelle-Aquitaine	34 (9,7%)	91 (6,6%)	7 (5,6%)
Occitanie	25 (7,2%)	117 (8,4%)	14 (11,1%)
Pays de la Loire	49 (14,0%)	58 (4,2%)	20 (15,9%)
Provence-Alpes-Côte d'Azur	19 (5,4%)	119 (8,6%)	5 (4,0%)
<b>Type d'hôpital</b>			
CHU	207 (59,3%)	683 (49,3%)	69 (54,8%)
CH	140 (40,1%)	558 (40,2%)	54 (42,8%)
Autre	2 (0,6%)	146 (10,5%)	3 (2,4%)
<b>Unité médicale</b>			
Réanimation adulte	274 (78,5%)	1102 (79,5%)	101 (80,2%)
Réanimation pédiatrique	75 (21,5%)	285 (20,5%)	25 (19,8%)

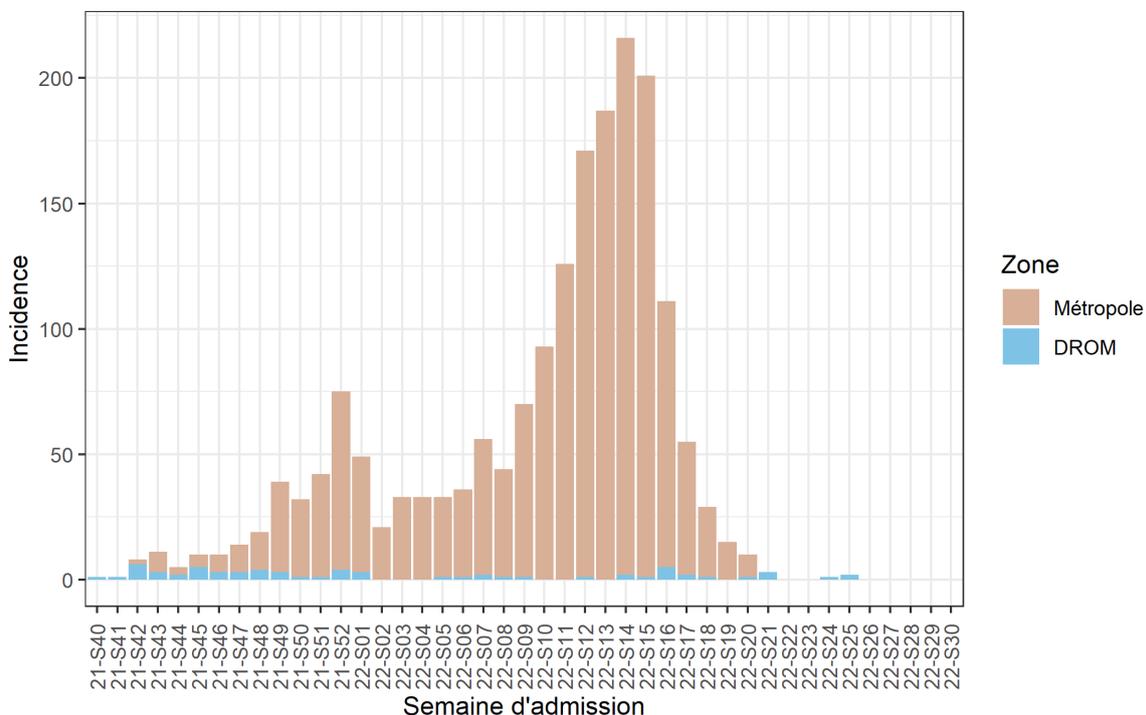
<sup>1</sup>n (%); Médiane (EI)

On peut remarquer que pour les variables âge et classe d'âge, les cas identifiés comme communs avaient une distribution qui semblait davantage se rapprocher de celle des cas uniquement dans le PMSI. Les distributions des effectifs étaient similaires pour les trois sous-groupes pour les variables « sexe » et « unité médicale ». Une plus grande hétérogénéité était observée dans la distribution des effectifs selon les régions. Les régions ayant le plus de cas communs étaient celles ayant déclaré le plus de cas dans le cadre de la surveillance SGS.

### 3.2.3 Description spatio-temporelle des cas graves de grippe admis en réanimation entre la métropole et les DROM

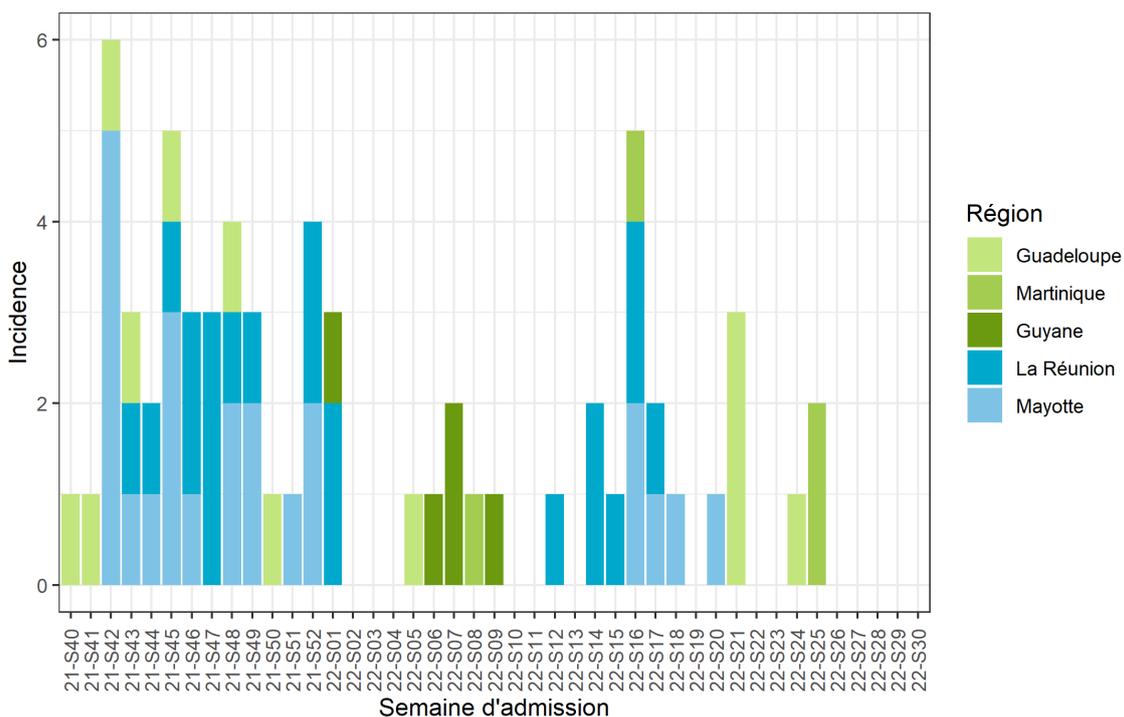
Lorsqu'on s'intéresse de plus près aux données issues des DROM, on remarque que celles-ci représentaient une proportion faible de cas, comme illustré dans la Figure 6. La distribution temporelle des données montre également que les cas des DROM ont surtout été déclarés au début et vers la fin de la période d'étude avec quelques déclarations sporadiques en milieu de la période (Fig. 6).

Figure 6 – Répartition de l'incidence des cas graves de grippe admis en réanimation par semaine d'admission entre la métropole et les DROM, Semaine 40 de 2021 à Semaine 30 de 2022



Les cas déclarés en début et en fin de saison provenaient essentiellement de la Guadeloupe, de la Réunion et de Mayotte. Les cas signalés en milieu de saison provenaient majoritairement de la Guyane (Fig.7).

Figure 7 – Incidence des cas graves de grippe admis en réanimation par semaine d'admission, DROM, Semaine 40 de 2021 à Semaine 30 de 2022

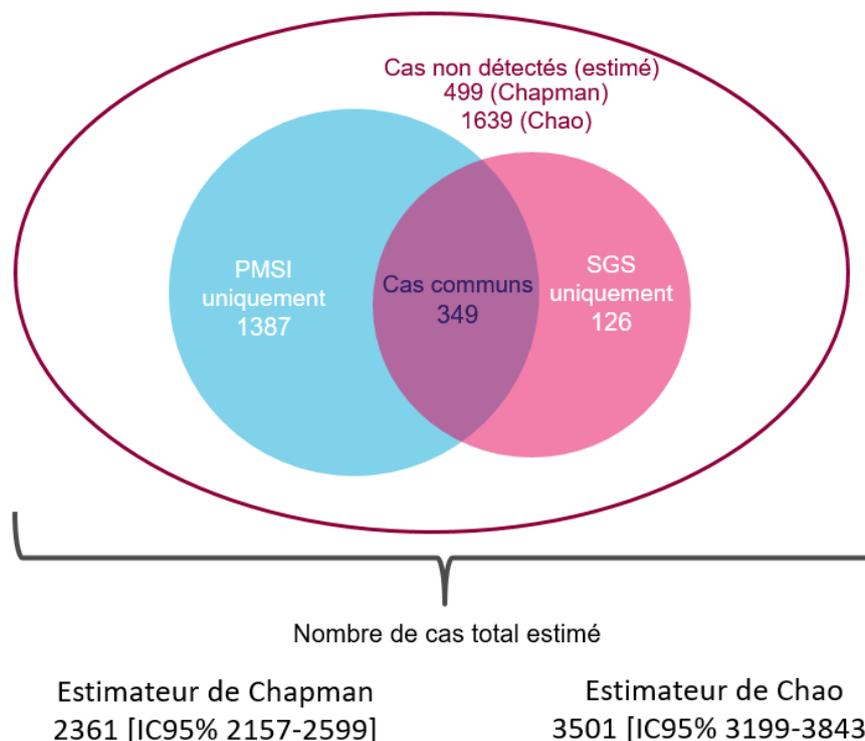


### 3.3 Estimation de l'exhaustivité des bases SGS et PMSI par capture-recapture

#### 3.3.1 Estimation du nombre de cas total au niveau national

Une fois que le nombre de cas communs et le nombre de cas spécifiques à chacune des deux bases a été déterminé, l'estimation du nombre de cas total a été calculée en utilisant l'estimateur de Chapman et de Chao. Ainsi, avec la méthode de Chapman, le nombre de cas total estimé est de 2361 (IC 95%, [2157-2599]), soit un nombre de cas non détectés par nos deux sources de 499 en moyenne. L'estimateur de Chao donne lui une valeur de N estimée à 3501 (IC95% [3199-3843]), soit environ 1,5 fois plus élevée qu'avec l'estimateur de Chapman, et un nombre de cas non détectés de 1639 en moyenne, soit environ trois fois plus élevé. Si on considère un cas idéal où tous les cas de grippe sévère admis en réanimation pourraient être retrouvés dans au moins une des deux bases, la valeur de N serait de 1862 (voir II.2.4.2). L'intervalle de plausibilité  $I_{Ch-BCS}$  était donc de [1862-3843].

Figure 8 – Diagramme de Venn des effectifs selon le résultat du croisement de la base PMSI et SGS pour les cas graves de grippe admis en réanimation avec estimation du nombre de cas total selon l'estimateur de Chapman et de Chao, analyse globale, France entière, Semaine 40 de 2021 à Semaine 30 de 2022



Dans le cas des analyses restreintes aux services sentinelles, le nombre de cas total estimé par les estimateurs de Chapman et Chao sont, respectivement, 1720 [1580-1888] et 2168 [1986-2380]. La valeur donnée par l'estimateur de Chao est environ 1,3 fois plus élevée que celle donnée par la méthode de Chapman. L'intervalle de plausibilité  $I_{Ch-BCS}$  était de [1391-2380].

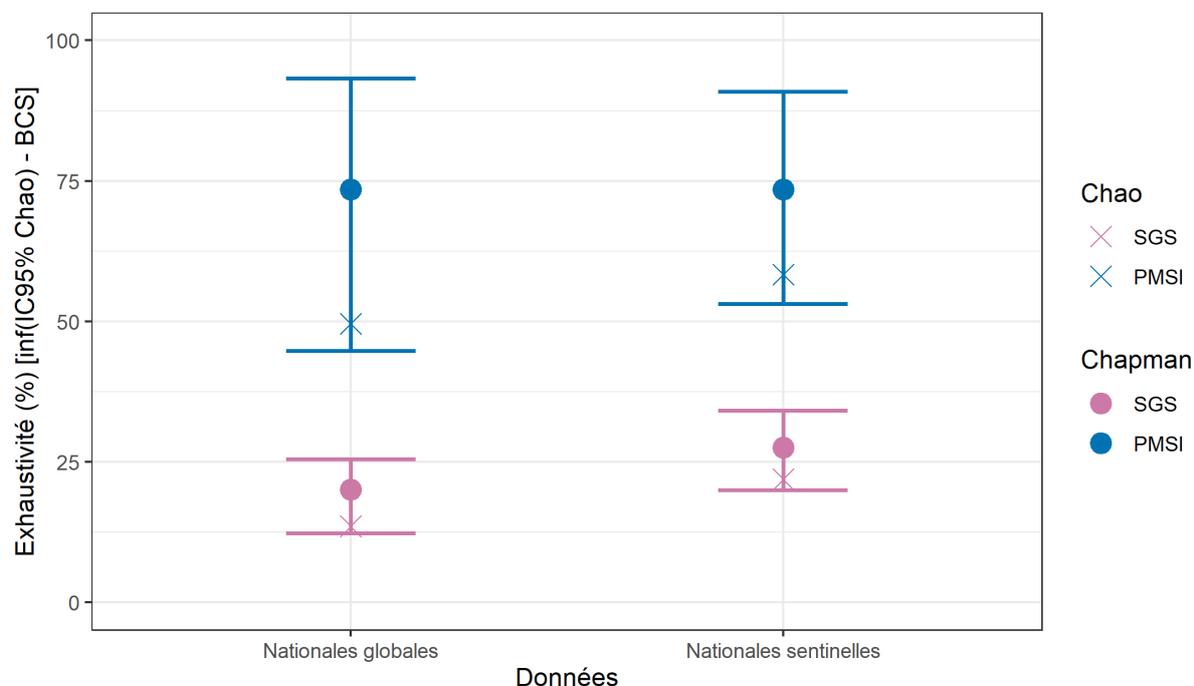
#### 3.3.2 Estimation de l'exhaustivité globale et restreinte aux services sentinelles

Le calcul de l'exhaustivité a été réalisé selon la méthode décrite précédemment, c'est-à-dire en divisant le nombre de cas dans chaque base par la valeur du nombre de cas estimé (voir II.2.3.3).

Au niveau national, l'exhaustivité de la surveillance SGS calculée à partir du nombre de cas estimé par la méthode Chapman était de 20,1% (IC 95%, [18,2-22,0%]) et l'exhaustivité du PMSI était de 73,5% [66,8-80,5%]. L'estimateur de Chao donnait une valeur de 13,6% (IC 95%, [12,3-14,8%]) pour

l'exhaustivité SGS et 49,6% (IC 95%, [44,7-54,3%]) pour l'exhaustivité PMSI. Si on considère le cas idéal (*best case scenario*), l'exhaustivité SGS serait de 25,5% et celle du PMSI serait de 93,2% (voir II.2.4.2.). L'intervalle de plausibilité  $I_{Ch-BCS}$  était donc de [12,3-25,5%] pour l'exhaustivité SGS et de [44,7-93,2%] pour l'exhaustivité PMSI (Fig.9).

Figure 9 – Estimation de l'exhaustivité par source au niveau national pour les analyses globales et restreintes aux services sentinelles, France entière, Semaine 40 de 2021 à Semaine 30 de 2022



En restreignant aux données de la métropole, l'exhaustivité de la surveillance SGS était de 20,1% (IC 95%, [18,2-22,0%]) avec l'estimateur de Chapman et de 13,3% (IC 95%, [12,1-14,6%]) avec l'estimateur de Chao. L'intervalle de plausibilité  $I_{Ch-BCS}$  était de [12,1-25,0%]. Pour le PMSI, l'exhaustivité était de, respectivement, 75,5% (IC 95%, [68,4-82,9%]) et 50,1% (IC 95%, [45,4-55,0%]). L'intervalle  $I_{Ch-BCS}$  était de [45,4-93,9%].

Un calcul de l'exhaustivité a également été réalisé en considérant uniquement les données des services considérés comme faisant partie du réseau sentinelles de la surveillance des cas graves (voir II.2.2.2). Ainsi, pour les analyses nationales restreintes aux services sentinelles, l'exhaustivité de la surveillance SGS était de 27,6% [25,2-30,1%] avec l'estimateur de Chapman et de 21,9% [20,0-23,8%] avec l'estimateur de Chao. Pour le PMSI, l'exhaustivité calculée selon les deux méthodes était, respectivement, 73,5% [67,0-80,0%] et 58,3% [53,1-63,5%]. L'intervalle  $I_{Ch-BCS}$  était donc de [20,0-34,1%] pour l'exhaustivité SGS et [53,1-90,9%] pour l'exhaustivité PMSI.

Au niveau métropolitain, l'exhaustivité de la base SGS et du PMSI pour les services sentinelles était légèrement plus élevée par rapport au niveau national mais sans différences majeures.

### 3.3.3 Analyses stratifiées de l'exhaustivité

Par la suite, l'exhaustivité a également été évaluée selon différentes variables de stratification : classe d'âge, région, période et type d'hôpital. La stratification par statut sentinelle n'a pas été réalisée étant donné que la strate des cas non sentinelle ne comportait aucun cas SGS et donc aucun cas commun. Dans cette situation, des doutes ont été soulevés quant à la fiabilité de l'estimateur de Chapman pour cette strate. Quant à l'estimateur de Chao, celui-ci était non défini (zéro au dénominateur).

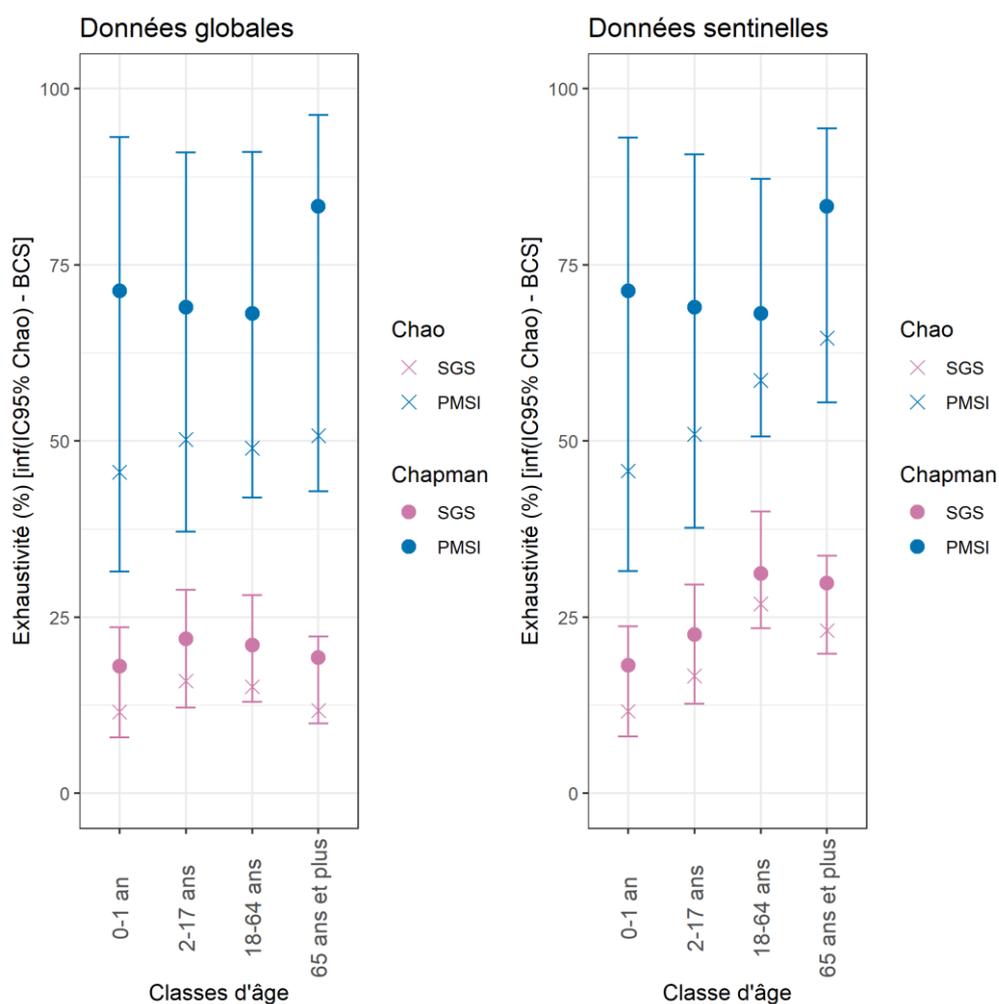
### 3.3.3.1 Stratification par classe d'âge

Une analyse de l'exhaustivité stratifiée par classe d'âge permettait d'observer une relative stabilité des valeurs de l'exhaustivité dans nos deux sources pour chaque catégorie d'âge dans le cadre des analyses globales à l'exception d'une exhaustivité en moyenne un peu plus élevée pour la catégorie des 65 ans et plus dans la base PMSI. (Fig.10).

En moyenne, les valeurs d'exhaustivité de la base SGS étaient comprises entre 18,0% et 21,9% avec l'estimateur de Chapman et entre 11,5% et 16% pour l'estimateur de Chao. Celles du PMSI étaient comprises entre 68,1% et 83,3% avec la méthode de Chapman et entre 45,6% et 50,8% avec l'estimateur de Chao.

Pour les analyses restreintes aux services sentinelles, les valeurs moyennes de l'exhaustivité SGS avaient augmenté pour la plupart des catégories et étaient comprises entre 18,2% et 31,2% pour l'estimateur de Chapman et entre 11,7% et 26,9% pour l'estimateur de Chao. La catégorie des 0-1 an était la seule à ne pas être concernée par cette augmentation de l'exhaustivité lors du passage aux analyses restreintes aux services sentinelles. L'exhaustivité PMSI ne présentait pas de variations notables par rapport aux résultats des analyses globales sauf au niveau des estimations de Chao qui augmentaient pour les catégories des 18-64 ans et des 65 ans et plus avec, respectivement, une valeur de 58,6% [50,7-66,6%] et 64,6% [55,4-74,3%].

Figure 10 – Estimation de l'exhaustivité par source et par classe d'âge pour les analyses globales et restreintes aux services sentinelles, France entière, Semaine 40 de 2021 à Semaine 30 de 2022

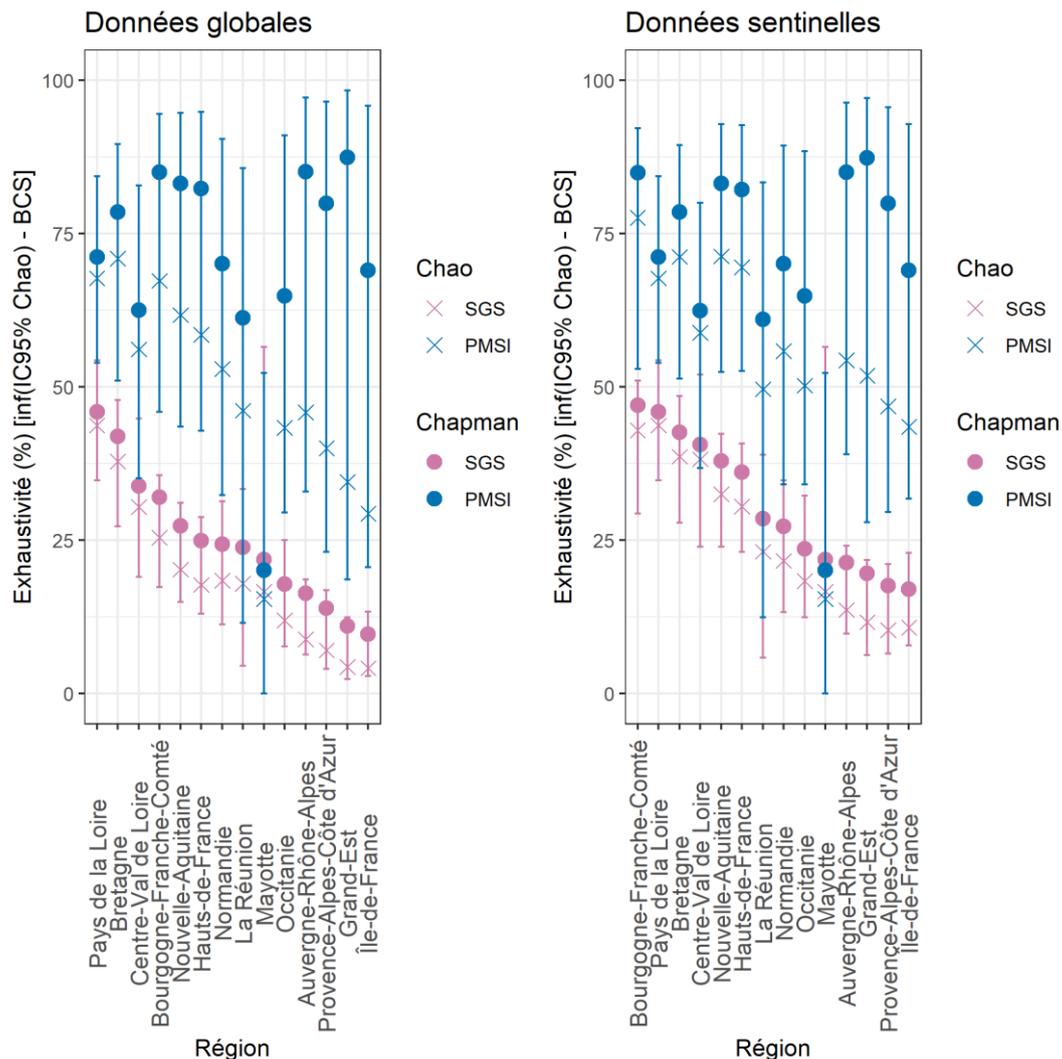


### 3.3.3.2 Stratification par région

L'analyse stratifiée par région montre une grande disparité au niveau des valeurs moyennes pour les deux bases. Au niveau de la base SGS, les valeurs moyennes d'exhaustivité étaient comprises entre 9,6% et 45,9% pour les analyses globales et entre 17,0% et 47,0% pour les analyses restreintes aux services sentinelles, avec l'estimateur de Chapman. En ayant recours à l'estimateur de Chao, les valeurs d'exhaustivité SGS étaient comprises entre 4,1% et 43,7% pour les analyses globales et entre 10,3% et 43,7% en restreignant aux services sentinelles. Pour la base PMSI, les valeurs moyennes d'exhaustivité étaient comprises entre 20,1% et 87,4% pour les données globales et entre 20,1% et 87,3% en restreignant aux services sentinelles, selon la méthode de Chapman. L'estimateur de Chao donnait des valeurs d'exhaustivité comprises entre 15,4% et 70,9% pour les analyses globales et entre 15,4% et 77,6% en restreignant aux services sentinelles (Fig.11). On remarque à nouveau une augmentation générale de l'exhaustivité lorsque seules les données issues des services sentinelles étaient prises en compte. Des cartes avec les exhaustivités par région sont en annexe (Annexe 6).

L'estimation du nombre de cas total obtenue à partir de la somme des estimations du nombre de cas par strate pour les analyses globales était de 2306 pour l'estimateur de Chapman et 4048 pour l'estimateur de Chao (Annexe 7).

Figure 11 – Estimation de l'exhaustivité par source et par région pour les analyses globales et restreintes aux services sentinelles, par ordre décroissant de la valeur moyenne de l'exhaustivité SGS calculée avec l'estimateur de Chapman, France entière, Semaine 40 de 2021 à Semaine 30 de 2022



La région avec l'exhaustivité moyenne SGS calculée par la méthode de Chapman la plus élevée était la région Pays de la Loire pour les analyses globales (45,9%) et Bourgogne-Franche-Comté (47%) pour les analyses restreintes aux services sentinelles. L'exhaustivité la plus faible pour les deux types de données était observée en Ile-de-France (9,6% pour les analyses globales et 17,0% pour les analyses sentinelles). Lorsque l'exhaustivité était calculée avec l'estimateur de Chao, c'était à nouveau la région Pays de la Loire qui présentait la plus grande exhaustivité mais cette fois-ci également pour les analyses sentinelles. C'est l'Ile-de-France et le Grand-Est qui présentaient les valeurs les plus faibles pour, respectivement, les analyses globales et sentinelles.

Au niveau du PMSI, l'exhaustivité moyenne calculée avec l'estimateur de Chapman la plus élevée était observée pour la région Grand-Est à la fois pour les analyses globales et sentinelles. L'exhaustivité la plus faible était observée à Mayotte. Cependant, le calcul de l'exhaustivité par l'estimateur de Chao montrait que la plus grande exhaustivité PMSI était observée pour la région Bretagne pour les analyses globales et Bourgogne-France-Comté pour les analyses sentinelles.

La figure 9 montre également que la valeur moyenne d'exhaustivité estimée par l'estimateur de Chao se rapprochait de l'estimation par la méthode de Chapman pour certaines régions, et ce dans les deux bases. Cela concernait notamment les régions où l'exhaustivité SGS était la plus élevée.

C'est lors de l'analyse descriptive des données stratifiées par région que nous avons observé des effectifs plutôt faibles pour certaines strates, ce qui nous avait poussé à privilégier l'utilisation de l'estimateur de Chapman au lieu de l'estimateur MRR. Une comparaison des résultats d'exhaustivité selon ces deux estimateurs a été réalisée (Tab.4).

Tableau 4 – Comparaison de l'estimateur de Chapman avec le *Mark Recapture Ratio* pour les exhaustivités SGS et PMSI globales par région

Estimateur	Chapman		MRR	
	Exhaustivité SGS <sup>1</sup>	Exhaustivité PMSI <sup>1</sup>	Exhaustivité SGS <sup>1</sup>	Exhaustivité PMSI <sup>1</sup>
Auvergne-Rhône-Alpes	16,3	85,1	16,2	84,8
Bourgogne-Franche-Comté	32,0	85,0	31,9	84,6
Bretagne	41,9	78,5	41,7	78,1
Centre-Val de Loire	33,8	62,5	33,3	61,5
Corse	N/A	N/A	N/A	N/A
Grand-Est	11,0	87,4	10,9	86,7
Guadeloupe	N/A	N/A	N/A	N/A
Guyane	N/A	N/A	N/A	N/A
Hauts-de-France	24,9	82,3	24,8	82,0
Île-de-France	9,6	69,0	9,5	68,5
La Réunion	23,8 [9,3-38,9]	61,2 [24,0-100]	22,2 [5,6-38,9]	57,1 [14,3-100]
Martinique	N/A	N/A	N/A	N/A
Mayotte	21,8 [7,2-44,3]	20,1 [6,6-40,9]	16,7 [0,0-41,7]	15,4 [0,0-38,5]
Normandie	24,3	70,1	24,0	69,2
Nouvelle-Aquitaine	27,3	83,2	27,2	82,9
Occitanie	17,8	64,8	17,6	64,1
Pays de la Loire	45,9	71,2	45,8	71,0
Provence-Alpes-Côte d'Azur	13,9	79,9	13,8	79,2

<sup>1</sup>Valeur moyenne (IC 95%)

N/A: non applicable ; exhaustivité non estimée (effectifs insuffisants)

Les estimations d'exhaustivité étaient globalement assez similaires pour les deux estimateurs. On observe néanmoins des différences plus marquées entre les estimateurs pour La Réunion et Mayotte.

### 3.3.3.3 Stratification par période de temps

Une analyse stratifiée par période de temps a également été menée. Deux périodes ont été considérées : une première s'étendant de la semaine 40 de 2021 à la semaine 08 de 2022 (période pré-épidémique) et une deuxième à partir de la semaine 09 de 2022 jusqu'à la semaine 30 de 2022 (période épidémique). L'objectif était de voir s'il y avait des différences de signalement (et donc d'exhaustivité) entre ces deux périodes.

Cette analyse a montré que l'exhaustivité variait faiblement entre ces deux périodes. Pour les données globales, l'exhaustivité SGS calculée avec l'estimateur de Chapman était de 20,0% (IC 95% [16,8-23,5%]) sur la première période et était de 20,2% [18,0-22,4%] sur la deuxième partie de la saison grippale. Avec l'estimateur de Chao, ces valeurs étaient respectivement de 13,9% [11,5-16,4%] et 13,4% [11,9-14,9%]. Les valeurs d'exhaustivité SGS semblaient donc globalement stables au cours du temps.

Les différences étaient un peu plus marquées au niveau du PMSI pour les valeurs moyennes avec une exhaustivité de 69,2% [58,1-81,0%] pour la première période et 75,6% [67,3-83,9%] sur la seconde période, avec l'estimateur de Chapman. Les résultats obtenus avec l'estimateur de Chao étaient, respectivement, 48,1% [39,9-56,4%] et 50,2% [44,7-55,8%].

Les mêmes observations peuvent être constatées au niveau des analyses sentinelles à l'exception que les exhaustivités SGS moyennes présentaient un écart un peu plus grand entre les deux périodes. Avec l'estimateur de Chapman, l'exhaustivité SGS était de 26,4% [22,2-30,9%] sur la première période et de 28,2 [25,3-31,2%] pour la deuxième période. Les exhaustivités calculées avec l'estimateur de Chao étaient, respectivement, 21,1% [17,7-24,7%] et 22,3% [19,9-24,6%]. (Annexe 4)

### 3.3.3.4 Stratification par type d'hôpital

Une stratification par type d'hôpital a été réalisée selon trois catégories : CHU, CH et Autre. Les établissements présents dans la catégorie « Autre » étaient principalement des cliniques et hôpitaux privés ainsi que des établissements du service de santé des armées, soit une certaine hétérogénéité dans le type de structures hospitalières.

Pour les analyses globales, les exhaustivités SGS calculées avec l'estimateur de Chapman des CHU et CH étaient du même ordre de grandeur avec 23,3% (IC 95% [20,6-26,0%]) pour les CHU et 20,1% [17,2-23,1%] pour les CH. Avec l'estimateur de Chao, on obtient 16,8% [14,8-18,8%] pour les CHU et 13,7% [11,7-15,7%] pour les CH. Il en était de même pour l'exhaustivité PMSI avec 75,1% [66,0-84,5%] pour les CHU et 72,3% [61,5-83,1%] pour les CH, avec l'estimateur de Chapman. L'estimateur de Chao donnait respectivement 54,2% [47,7-60,5%] et 49,1% [41,8-56,5%]. Les exhaustivités SGS et PMSI pour la catégorie « Autre » étaient très inférieures à celles des CHU et CH.

En restreignant l'analyse aux services sentinelles, l'exhaustivité SGS et PMSI des CHU ne présentait pas de grandes différences par rapport à l'analyse globale avec une exhaustivité SGS de 24,4% [21,6-27,3%] et une exhaustivité PMSI de 75,1% [66,4-84,1%], pour l'estimateur de Chapman. Les estimations de Chao donnaient 18,1% [16,0-20,3%] pour l'exhaustivité SGS et 55,6% [49,1-62,0%] pour l'exhaustivité PMSI des CHU. Cependant, une forte augmentation de l'exhaustivité SGS était observée au niveau des CH avec une valeur de 35,5% [31,0-40,3%], assez proche des résultats obtenus avec l'estimateur de Chao qui donnaient 31,3% [27,3-35,6%]. L'exhaustivité PMSI des CHU était quasiment inchangée. Celle des CH était inchangée pour l'estimation avec Chapman mais avait une valeur de 63,8% [55,1-72,4%] avec l'estimateur de Chao. Une augmentation de l'exhaustivité SGS et surtout de l'exhaustivité PMSI était observée pour la catégorie « Autre », pour les deux estimateurs (Annexe 5).

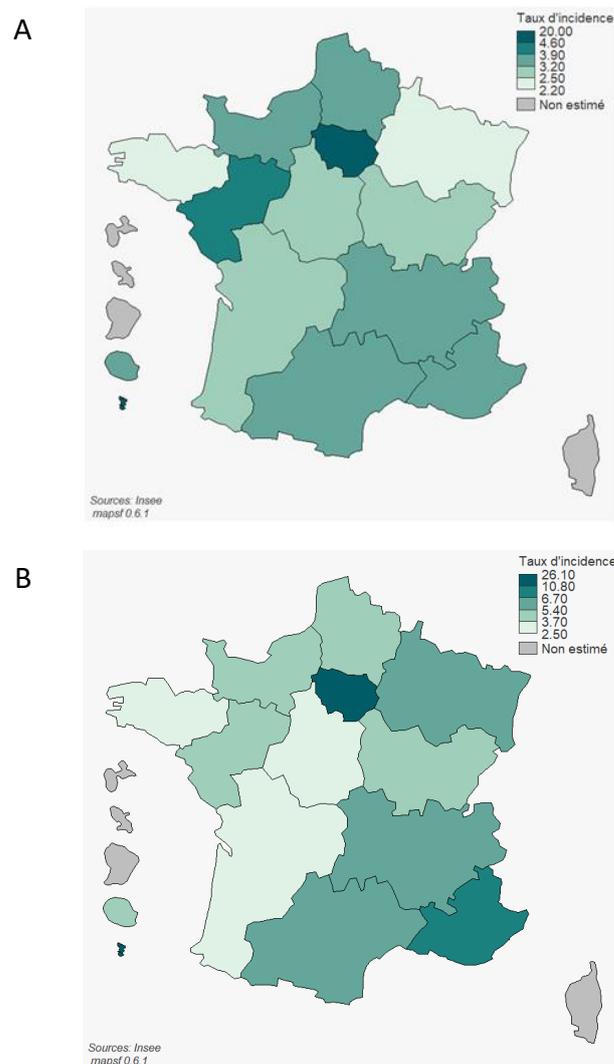
### 3.3.4 Estimation du taux d'incidence des cas graves de grippe admis en réanimation

L'estimation du nombre de cas total de cas de grippe sévère admis en réanimation pour chaque région permet également d'estimer le taux d'incidence pour cent mille habitants des cas de grippe sévères à l'échelle nationale et régionale.

Au niveau national, l'incidence des cas graves de grippe admis en réanimation était estimée à 3,5/100 000 habitants [3,2-3,8] avec l'estimateur de Chapman et 5,2/100 000 habitants [4,7-5,7] avec l'estimateur de Chao. L'intervalle de plausibilité était de [2,7-5,7]. Les données de population des régions où l'exhaustivité n'a pas été estimée ont été prises en compte pour le calcul des incidences nationales.

Les régions avec le taux d'incidence le plus élevé étaient Mayotte avec un taux d'incidence moyen de 20/100 000 habitants avec l'estimateur de Chapman et 26,1 avec l'estimateur de Chao et l'Île-de-France avec un taux de 4,6 (10,8 avec l'estimateur de Chao). Les régions avec les taux d'incidence les plus faibles étaient la Bretagne (2,2/100 000 habitants) et le Grand-Est (2,4) avec l'estimateur de Chapman et la Bretagne (2,5), la Nouvelle-Aquitaine (3,3) et le Centre-Val de Loire (3,3) avec l'estimateur de Chao (Fig.12).

Figure 12 – Carte du taux d'incidence pour 100 000 habitants des cas graves de grippe admis en réanimation estimé par région, à partir des valeurs moyennes de l'estimateur de Chapman (A) ou de Chao (B) issues de l'analyse globale



## 4. Discussion

L'étude Ecsir réalisée sur la saison grippale 2021-2022 est la première analyse d'exhaustivité du système de surveillance des gripes sévères depuis la pandémie de COVID-19. Cette dernière a eu un impact majeur sur la surveillance avec par exemple le passage en surveillance conjointe COVID/grippe et la saturation des services de réanimation. Une évaluation du système de surveillance était donc nécessaire pour en mesurer la performance après plusieurs années de pandémie COVID.

Au niveau national, l'exhaustivité SGS était vraisemblablement comprise sur l'intervalle [12,3-25,5%] pour les analyses globales et [20,0-34,1%] pour les analyses restreintes aux données sentinelles. La comparabilité avec les résultats des études précédentes était limitée étant donné que nous avons eu recours à une méthodologie différente. Une différence majeure provenait du fait d'avoir utilisé la variable date d'admission en réanimation dans les deux bases pour le croisement des sources à la place de la date d'hospitalisation pour le PMSI et la date d'admission en réanimation pour la base SGS. Cependant, les valeurs obtenues semblaient tout de même très faibles par rapport aux exhaustivités SGS globales estimées de 47% (IC 95% [47-48%]) pour la période 2009-2013 (Loury *et al.*, 2017), 46,1% [45,5-46,8%] pour la période 2017-2018 (Pivette *et al.*, 2019) et 36,3% [35,6-37,0%] pour la période 2018-2019 (données non publiées). Pour les périodes 2009-2013 et 2017-2018, la surveillance SGS était encore à visée exhaustive. Les chiffres sur la saison 2018-2019 sont les premiers depuis le passage à une surveillance sentinelle. La comparaison des résultats ne peut donc se faire qu'avec les données de 2018-2019. Une nouvelle analyse reprenant les mêmes modalités que la dernière étude Ecsir sera réalisée ultérieurement pour que la comparabilité soit pertinente.

Les analyses restreintes au niveau métropolitain n'ont pas vraiment montré de différences par rapport aux analyses nationales. Cela peut s'expliquer par la faible part des cas déclarés par les DROM par rapport aux cas identifiés en métropole (Fig.6).

Les exhaustivités estimées au niveau régional étaient très disparates. Les régions avec la meilleure exhaustivité SGS au niveau des analyses globales étaient les régions Pays de la Loire (45,9% [36,7-55,2] avec l'estimateur de Chapman, 43,7% [34,7-52,6%] avec l'estimateur de Chao) et Bretagne (41,9% [30,5-53,3%] avec l'estimateur de Chapman, 37,8% [27,2-48,4%] avec l'estimateur de Chao). L'exhaustivité SGS la plus faible était observée dans la région Ile-de-France (9,6% [6,8-12,6%] avec l'estimateur de Chapman, 4,1% [2,8-5,4%] avec l'estimateur de Chao). Cependant, ces disparités régionales étaient attendues étant donné que la sélection des services participant à la surveillance a été réalisée par chaque CR et que la proportion de services suivis est variable d'une région à une autre.

Au niveau des analyses restreintes aux services sentinelles, les régions avec la meilleure exhaustivité étaient la région Pays de la Loire (45,9% [36,7-55,2%] avec l'estimateur de Chapman, 43,7% [34,7-52,6%] avec l'estimateur de Chao) et la région Bourgogne-Franche-Comté (47,0% [32,5-55,3%] avec l'estimateur de Chapman, 42,9% [29,3-50,7%] avec l'estimateur de Chao). L'Ile-de-France était ici encore la région avec la plus faible exhaustivité SGS (17,0% [12,1-22,0%] avec l'estimateur de Chapman, 10,7% [7,8-13,9%] avec l'estimateur de Chao), suivie de près par la région Provence-Alpes-Côte d'Azur (17,6% [11,4-22,0%] avec l'estimateur de Chapman, 10,3% [6,5-13,0%] avec l'estimateur de Chao). Ces différences au niveau de l'analyse restreinte aux services sentinelles pouvaient résulter de différences de participation des services à la surveillance (variabilité de la participation selon les hôpitaux concernés, différences des modalités d'animation du réseau par les CR entre les régions etc.).

Les taux d'incidence globaux estimés différaient légèrement selon les estimateurs. Cependant, dans les deux cas, les régions où le taux d'incidence était le plus élevé étaient Mayotte (taux d'incidence moyen de 20/100 000 habitants avec l'estimateur de Chapman et 26,1 avec l'estimateur de Chao) et l'Ile-de-France (taux de 4,6 avec l'estimateur de Chapman, 10,8 avec l'estimateur de Chao) (Fig. 12). La Bretagne était la région avec le plus faible taux d'incidence (2,2/100 000 habitants avec l'estimateur de Chapman, 2,5 avec l'estimateur de Chao). Ces différences d'incidence entre les régions avaient déjà été observées dans les précédentes études (Loury *et al.*, 2017). Les hypothèses envisagées

pour les expliquer étaient des variations d'incidence de la grippe sur le territoire, les caractéristiques de la population dans chaque région (âge, catégories sociales etc.) qui peuvent influencer sur le recours aux soins ainsi que sur l'admission en réanimation (critères d'âge), le climat (influence sur la circulation du virus), le nombre de lits de réanimation dans les hôpitaux etc.

Alors que les analyses globales permettent surtout d'estimer la proportion de cas signalés sur l'ensemble des cas en France, la performance du réseau de surveillance SGS est, elle, estimée grâce aux analyses restreintes aux services sentinelles. On observe globalement que les résultats d'exhaustivité SGS étaient plus élevés lorsque l'analyse était restreinte aux données sentinelles, ce qui était attendu étant donné que les cas SGS représentent une proportion importante des cas sentinelles. Une variabilité importante de l'exhaustivité a été observée selon les régions dans les deux niveaux d'analyse (Fig. 11). La stratification par classe d'âge a également montré que le réseau de surveillance des gripes sévères présentait en moyenne une exhaustivité plus élevée pour la catégorie des 18-64 ans ainsi que pour les 75 ans et plus. L'exhaustivité SGS était également en moyenne meilleure pour les hôpitaux de type CH. Cela peut s'expliquer par le fait que de nombreux services ayant été retenus lors du passage en surveillance sentinelle étaient des services de Centres Hospitaliers.

Nos analyses ont mis en évidence un niveau de signalement des cas plutôt constant au cours du temps (pas de différence d'exhaustivité entre les périodes pré-épidémiques et épidémiques) (Annexe 5), avec une courbe épidémique des cas SGS quasi identique dans son allure avec celle du PMSI (Fig. 3). Les valeurs d'exhaustivité SGS restaient malgré tout plutôt faibles, notamment lorsque l'analyse était restreinte aux services sentinelles. Cela pourrait être expliqué par l'impact important de la pandémie de la COVID-19 sur le signalement des cas graves de grippe admis en réanimation. La période de la saison grippale 2021-2022 correspondait également à une période de nouvelle vague de la pandémie. Par ailleurs, l'hypothèse d'un impact du COVID sur la circulation des virus grippaux ainsi que sur la surveillance de la grippe a déjà été suggérée dans le bilan de la saison grippale 2021-2022 de Santé publique France (Equipes de surveillance de la grippe, 2022). La surveillance de la grippe étant conjointe avec celle de la COVID-19, il est possible que les services de réanimation, qui étaient déjà surchargés, aient préféré privilégier le signalement des cas graves de COVID-19 plutôt que les cas graves de grippe. Cependant, une étude (non publiée) réalisée sur l'exhaustivité de la surveillance des cas graves de COVID (période 2020-2021) montrait une exhaustivité plutôt faible de cette surveillance, ce qui est en défaveur de cette hypothèse. Il est alors probable que la surcharge générale des services de réanimation ait amené à une faible participation, que ce soit pour la surveillance des cas graves de grippe ou de COVID.

Si on observe les résultats des exhaustivités PMSI, les estimations données par l'estimateur de Chapman (73,5% [66,8-80,5%] au niveau national global) étaient assez en accord avec celles des études Ecsir précédentes. Cependant, les résultats donnés par l'estimateur de Chao étaient bien inférieurs (49,6% (IC 95%, [44,7-54,3%])). Ces résultats étaient assez inattendus étant donné qu'une stabilité de l'exhaustivité PMSI avait été observée depuis la première étude Ecsir. Nous pouvons supposer que la double épidémie grippe-COVID ainsi que le contexte pandémique a pu amener davantage de cas de grippe à être signalés comme des cas COVID dans la base PMSI dans le cadre d'une co-infection grippe/COVID, ce qui pourrait amener à une diminution de l'exhaustivité PMSI. Il pourrait être également intéressant de voir si ce contexte épidémique ajouté à la surcharge de l'activité hospitalière n'a pas pu amener les saisies PMSI à davantage se reposer sur les données de réanimation. Une description détaillée du parcours du recueil des données pourrait être fortement utile à l'interprétation des résultats d'exhaustivité.

La comparaison des résultats d'exhaustivité obtenus par le MRR et l'estimateur de Chapman pour les régions montrait quelques différences au niveau de La Réunion et Mayotte. Les intervalles de confiance bootstrap à 95% associés à ces valeurs moyennes étaient très larges et se recoupaient. Les estimations de l'IC étaient donc peu informatives, notamment pour le PMSI. La présence d'un nombre de cas faibles pour ces régions était probablement à l'origine des différences observées pour les valeurs moyennes et d'un manque de puissance de calcul amenant à des grands IC.

Il n'est pas forcément aisé de définir à partir de quel seuil des effectifs peuvent être considérés comme faibles. C'est la raison pour laquelle l'utilisation de l'estimateur de Chapman serait sûrement à privilégier systématiquement par rapport au MRR lorsque les conditions d'application sont vérifiées. Les IC obtenues restaient parfois toujours peu informatives mais étaient un peu plus précises que celles obtenues par le MRR. Cela pose également la question de l'utilisation de manière plus générale d'une forme de l'estimateur de Chao moins biaisée pour les petits effectifs (Annexe 8). De plus, l'estimateur de Chao n'est pas défini lorsque le nombre de cas commun est nul, ce qui n'est pas le cas de la version de l'estimateur adaptée aux faibles effectifs. L'application de cet estimateur corrigé pour les données régionales montrait que les valeurs étaient globalement similaires à celles données par la formule de Chao initiale sauf pour Mayotte et La Réunion où les estimations d'exhaustivité moyenne étaient plus élevées.

Des incertitudes persistaient néanmoins dans nos analyses et sont à prendre en considération pour l'interprétation des résultats ainsi que pour évaluer leur fiabilité. Une attention particulière est à porter sur les points suivants.

La sélection des cas déclarés par des services de réanimation sentinelles dans nos deux bases comportait quelques incertitudes. En effet, le statut sentinelle d'un service était disponible uniquement dans la base des services. Celle-ci a donc été fusionnée avec la base SGS grâce à la variable identifiant un service et avec la base PMSI grâce à la variable FINESS géographique (identifiant de service absent dans le PMSI). Cependant, plusieurs services d'un même hôpital peuvent avoir le même numéro FINESS alors que ce sont des services différents et que tous ne sont pas considérés comme sentinelle. Étant donné qu'il était impossible de retrouver le service ayant déclaré un cas à partir de nos données PMSI, nous avons considéré que pour un numéro FINESS donné, si au moins un des services associés à ce numéro était sentinelle, alors le cas relié à ce numéro FINESS était considéré comme déclaré par un service sentinelle. Une vérification manuelle du nombre de services concernés par ces ambiguïtés d'attribution du statut sentinelle nous a permis de voir que cela représentait seulement un faible pourcentage de cas PMSI (environ 5%). Nous pouvons donc supposer que les cas sentinelles étaient assez bien identifiés dans notre étude. Cependant, il pourrait être intéressant de voir s'il serait possible à l'avenir d'obtenir des données permettant de distinguer les services dans le PMSI pour gagner en précision.

La méthode de capture-recapture s'applique selon des conditions assez strictes et qui varient selon les estimateurs utilisés. Cela concerne notamment les hypothèses d'homogénéité de capture et d'indépendance des sources. Ces dernières sont indispensables à l'utilisation du MRR et de l'estimateur de Chapman : en cas de non-respect de ces hypothèses, les estimations sont biaisées et le nombre total de cas est surestimé ou sous-estimé (Hook et Regal, 1995 ; Gallay et al., 2002). L'estimateur de Chao constitue une alternative plus fiable dans cette situation étant donné qu'il est moins sensible à ces hypothèses (il prend en compte l'hétérogénéité de capture) (Chao, 1987 ; Chao, 1989 ; Brittain et Böhning, 2009). Globalement, dans nos analyses, les valeurs du nombre de cas total estimées par l'estimateur de Chao sont supérieures de façon significative par rapport aux estimations de Chapman (non recoupement des IC 95%). Cela se traduit par des exhaustivités plus basses par rapport à la méthode de Chapman. La présence d'une dépendance entre les sources ou d'une hétérogénéité de capture pourrait expliquer les différences observées entre les deux estimations.

Nous pouvons alors supposer qu'il existe probablement un lien de dépendance positive entre la base SGS et la base PMSI amenant à une sous-estimation du nombre de cas total et donc une surestimation de l'exhaustivité par l'estimateur de Chapman. Les résultats donnés par l'estimateur de Chao sont donc plus fiables et moins biaisés dans cette situation. L'application de cet estimateur à des données réelles ou des simulations par Chao (1987) ou Brittain et Böhning (2009) ont montré des bonnes performances de cet estimateur pour approximer la vraie valeur de N en comparaison avec d'autres estimateurs, même si l'estimateur donne un minorant de N.

Il est établi que l'homogénéité des sources implique l'indépendance (IWGDMF, 1995). Ceci équivaut à dire que la dépendance implique l'hétérogénéité (contraposée). Nos résultats ont montré des disparités dans la distribution des effectifs par région et par type d'hôpital. Un objectif des analyses stratifiées a été de prendre en compte l'effet d'éventuels facteurs d'hétérogénéité et donc d'estimer  $N$  sur des strates homogènes. Au niveau des analyses globales et avec l'estimateur de Chapman, les valeurs de  $N$  estimées par stratification étaient toutes contenues dans l'intervalle de confiance de  $\hat{N}$  [2157-2599] estimé sans stratification. Ceci pourrait témoigner vraisemblablement d'une absence d'ajustement de la valeur de  $\hat{N}$  selon les strates considérées et donc d'une relative homogénéité. Cependant, la valeur estimée à partir de l'analyse stratifiée par région ne prend pas en compte les régions exclues pour cause de faible effectif (voir II.2.5), ce qui limite l'interprétation de cette valeur. En considérant l'estimateur de Chao,  $\hat{N}$  était en dehors de l'intervalle [3199-3843] pour la stratification par région ( $\hat{N} = 4048$ ) et par type d'hôpital ( $\hat{N} = 7284$ ). Pour la stratification par type d'hôpital, il était possible que ce résultat soit dû à un manque de puissance de calcul dû à des effectifs particulièrement faibles dans certaines strates. Le calcul des valeurs estimées de  $N$  en utilisant une version modifiée de l'estimateur de Chao adaptée pour les faibles effectifs (Chao, 1989 ; Rivest et Baillargeon, 2007) donnait, par somme des strates de la variable type d'hôpital,  $\hat{N} = 5879$  (formule en Annexe 8). Cette valeur restait néanmoins encore en dehors de l'intervalle [3199-3843] alors que la valeur calculée avec la méthode de Chapman était, elle, bien comprise dans l'IC associé.

Ainsi, les différences d'estimation entre les deux estimateurs, de même que les différences dans les résultats des estimations de  $N$  stratifiées, semblaient plutôt en faveur de la présence d'une hétérogénéité de capture et d'une dépendance positive pour les sources étudiées. On remarque au niveau de l'analyse stratifiée par région que la valeur d'exhaustivité calculée par l'estimateur de Chao pour les régions Pays de la Loire, Centre-Val de Loire et Bretagne se rapprochait de celle de l'estimateur de Chapman pour les deux bases (Fig. 9). Les deux estimateurs peuvent avoir des valeurs similaires lorsque  $N_{SGS} = N_{PMSI}$  (Brittain et Böhning, 2009). Ce n'était cependant pas le cas ici. Nous pouvons alors supposer que les données issues de ces régions se rapprochaient des conditions où il y a homogénéité de capture et indépendance des sources. Ces régions sont d'ailleurs celles où les meilleures exhaustivités SGS sont observées. Nous pouvons supposer qu'une meilleure déclaration des cas SGS pourrait permettre d'avoir des effectifs SGS plus représentatifs de la population d'étude et se rapprocher d'une situation d'homogénéité de capture. Ce meilleur signalement des cas pourrait aussi être lié aux modalités d'animation du réseau de surveillance dans ces régions (relances plus régulières par les CR, meilleure adhésion des hôpitaux au dispositif de surveillance etc.).

En outre, les estimations de  $N$  par l'estimateur de Chao dans le cadre des analyses stratifiées restreintes aux services sentinelles avaient des valeurs beaucoup plus proches des valeurs estimées selon la méthode de Chapman (Annexe 7). De plus, toutes les valeurs de  $\hat{N}$  obtenues par stratification pour les deux estimateurs étaient comprises dans l'IC de  $\hat{N}$  sans stratification. Nous pouvons supposer que les données issues des services sentinelles se rapprocheraient davantage des conditions où les hypothèses d'homogénéité et d'indépendance sont vérifiées. En effet, le statut sentinelle ou non d'un service de réanimation impacte fortement sa probabilité de capture. Les cas admis dans des services non sentinelles avaient une probabilité nulle d'être détectés par la surveillance SGS. En restreignant l'analyse aux services sentinelles, ce facteur important d'hétérogénéité est supprimé.

Comme évoqué précédemment, au niveau des analyses globales, les cas non admis dans un service de réanimation sentinelle avaient une probabilité de détection par la surveillance SGS nulle. Nos conditions d'application ne semblaient pas mentionner cette situation comme disqualifiante pour l'utilisation de la capture-recapture. Des simulations visant à tester les performances de l'estimateur de Chao semblaient inclure des situations où les probabilités de capture pouvaient être nulles (Chao, 1987). Nous avons donc considéré que la capture-recapture était applicable.

Une description plus détaillée des modalités de recueil des données pour les bases SGS et PMSI en interrogeant les acteurs concernés pourrait être hautement profitable pour les prochaines études Ecsir pour permettre l'identification d'éventuels liens de dépendance. Il pourrait également être intéressant

de mener une nouvelle enquête sur les modalités d'animation des services de réanimation comme celle réalisée dans le cadre de la première étude Ecsir en 2017 (Loury *et al.*, 2017). Cela permettrait notamment de voir des éventuelles différences entre régions et des changements de pratique depuis la dernière étude, donnant alors éventuellement des éléments pour mieux interpréter nos résultats ainsi que pour améliorer la surveillance.

En outre, une hypothèse peu discutée concernant l'estimateur de Chao est que même si chaque individu a une probabilité de capture différente (prise en compte de l'hétérogénéité), cette probabilité est constante pour toutes les captures (Chao, 1987 ; Chao, 1989). Il est peu probable que cette hypothèse soit vérifiée dans notre étude étant donné que le recueil de données PMSI est systématique et que la surveillance SGS est basée sur le volontariat et concerne un nombre restreint de services dits sentinelles. Van Hest (2008) suggère que le non-respect de cette hypothèse d'équiprobabilité amène l'estimateur de Chao à surestimer le nombre de cas  $N$  (et donc à sous-estimer l'exhaustivité de  $N$ ), du moins pour des petites variations des probabilités individuelles entre les sources. Cela ferait donc de l'exhaustivité estimée par l'estimateur de Chao une borne inférieure de l'exhaustivité réelle. Dans cette situation, l'intervalle  $I_{Ch-BCS}$  serait plus précis.

Le choix d'une présentation des résultats également sous forme d'un intervalle de plausibilité, noté ici  $I_{Ch-BCS}$ , provenait des incertitudes concernant la précision de nos estimations et plus généralement dans le cadre d'une capture-recapture appliquée dans un contexte d'étude épidémiologique. La méthode utilisée par Hassel *et al.* (2011) a été reprise dans cette étude. L'application de l'intervalle  $I_{Ch-BCS}$  semblait bien adaptée à la base SGS mais beaucoup moins pour la base PMSI où les intervalles étaient très larges. Ceci est notamment dû au fait que l'exhaustivité PMSI donnée par l'estimateur de Chao, et donc la borne de son IC 95%, avait une valeur beaucoup plus faible que la valeur BCS (cas idéal). Cela pourrait s'expliquer par l'impact de la pandémie COVID sur le codage PMSI comme discuté précédemment. Cependant, la largeur des intervalles pour le PMSI n'est pas forcément gênante considéré que le système que l'on cherche avant tout à évaluer est la surveillance SGS. Néanmoins, dans une situation où il y a hétérogénéité et dépendance positive, les estimations par la méthode de Chapman sous-estiment le vrai nombre de cas et donc surestiment l'exhaustivité associée. Par conséquent, une manière d'obtenir des intervalles plus restreints pourrait être d'utiliser l'estimation de Chapman ou la borne supérieure de son IC à la place de la valeur obtenue pour le *best case scenario*, si sa valeur est effectivement inférieure au BCS.

Par ailleurs, comme rappelé dans Hook et Regal (1995), si tous ces éléments peuvent nous faire douter de la précision ou de la fiabilité des résultats et incitent à la prudence quant à l'interprétation des estimations brutes, les données présentées ici sont néanmoins importantes pour suivre les tendances d'une année à l'autre et identifier des leviers d'améliorations. Ces informations sont utiles pour l'évaluation de la surveillance, notamment pour les CR. Ces résultats d'exhaustivité leur permettent d'adapter leur relance ou les modalités de rétro-information pour améliorer la surveillance, ainsi que de nuancer les chiffres de la surveillance lors de la rédaction de bilans. Tous ces éléments sont en faveur d'une analyse d'exhaustivité régulière par le biais des études Ecsir.

### III - Conclusion

Notre étude est la première étude Ecsir depuis le début de la pandémie de COVID-19. Il est sans conteste que cette dernière a eu un impact important sur le système hospitalier (surcharge de travail pour les services de réanimation), le système de surveillance des gripes sévères (recueil et transmission des données, mise en place d'une surveillance conjointe) se traduisant par une forte baisse de l'exhaustivité de cette surveillance, et même sur la circulation des virus grippaux. Au vu des données recueillies dans cette étude, il nous paraît pertinent de réviser la composition et les modalités d'animation du réseau de surveillance SGS ainsi que les modalités de recueil des données. Une nouvelle sélection des services sentinelles pourrait ainsi être à considérer, en privilégiant les services avec les meilleurs taux de participation et de complétude des données. Par ailleurs, une dématérialisation du signalement pour les services sentinelles est envisagée à court terme, ce qui permettrait de simplifier la remontée des données, limiter les erreurs de saisie (valeurs de remplissage prédéfinies, harmonisation des notations et des formats, champs obligatoires ou non) et diminuer la charge de travail représentée par la saisie des fiches cliniques en ligne pour les cellules régionales de Santé publique France, afin de dégager plus de temps consacré à l'animation du réseau.

Cependant, l'idéal resterait une collecte des données directement depuis les logiciels métiers, en routine et de manière automatisée, à l'instar par exemple de ce qui a été développé aux urgences dans le cadre de la surveillance syndromique (réseau OSCOUR®). Le développement de systèmes d'information connectés sur les registres médicaux constitue ainsi un enjeu majeur pour la surveillance des pathologies infectieuses, particulièrement celles à potentiel pandémique comme la grippe, dans les années à venir.

## Références bibliographiques

- Agence Technique de l'Information sur l'Hospitalisation (ATIH). 2022a. Classification statistique internationale des maladies et des problèmes de santé connexes. CIM-10 FR à usage PMSI. Volume 1.
- Agence Technique de l'Information sur l'Hospitalisation (ATIH). 2022b. Guide méthodologique de production des informations relatives à l'activité médicale et sa facturation en médecine, chirurgie, obstétrique et odontologie.
- BRITAIN, S. & BÖHNING, D. 2009. Estimators in capture–recapture studies with two sources. *ASTA Advances in Statistical Analysis*, 93, 23-47.
- CAMPESE, C., JARRAUD, S., SOMMEN, C., MAINE, C. & CHE, D. 2013. Legionnaires' disease in France: sensitivity of the mandatory notification has improved over the last decade. *Epidemiol Infect*, 141, 2644-9.
- CHAO, A. 1987. Estimating the Population Size for Capture-Recapture Data with Unequal Catchability. *Biometrics* 43, 783-791.
- CHAO, A. 1989. Estimating Population Size for Sparse Data in Capture-Recapture Experiments. *Biometrics*, 45, 427.
- CHAO, A., TSAY, P. K., LIN, S. H., SHAU, W. Y. & CHAO, D. Y. 2001. The applications of capture-recapture models to epidemiological data. *Stat Med*, 20, 3123-57.
- CHAO, A. & COLWELL, R. K. 2017. Thirty years of progeny from Chao's inequality: Estimating and comparing richness with incidence data and incomplete sampling. *Sort-statistics and Operations Research Transactions*, 1, 3-54.
- CHAPMAN DG. 1951. Some properties of the hypergeometric distribution with applications to the zoological sample censuses. *Uni Calif Public Stat* 1951; 1: 1060-7
- Centre d'épidémiologie sur les causes médicales de décès (CépiDc). *Les statistiques sur les causes médicales de décès de A à Z*. [en ligne]. Mis à jour le 11 août 2022. [<https://www.cepidc.inserm.fr/qui-sommes-nous/les-statistiques-sur-les-causes-medicales-de-deces-de-z>] (consulté le 30 mai 2023).
- Centers for Disease Control and Prevention (CDC), a. *1918 Pandemic (H1N1 virus)* [en ligne]. Mis à jour le 20 mars 2019. [<https://www.cdc.gov/flu/pandemic-resources/1918-pandemic-h1n1.html>] (consulté le 9 février 2023)
- Centers for Disease Control and Prevention (CDC), b. *Key Facts About Influenza (Flu)* [en ligne]. Mis à jour le 24 octobre 2022. [<https://www.cdc.gov/flu/about/keyfacts.htm>] (consulté le 21 février 2022)
- Centers for Disease Control and Prevention (CDC), c. *Types of Influenza Viruses* [en ligne]. Mis à jour le 2 décembre 2022. [<https://www.cdc.gov/flu/about/viruses/types.htm>] (consulté le 21 février 2022)
- Centers for Disease Control and Prevention (CDC), d. *How Flu Viruses Can Change : « Drift » and « Shift »* [en ligne]. Mis à jour le 12 décembre 2022. [<https://www.cdc.gov/flu/about/viruses/change.htm>] (consulté le 21 février 2023)
- Centers for Disease Control and Prevention (CDC), e. *What You Should Know About Flu Antiviral Drugs* [en ligne]. Mis à jour le 15 décembre 2022. [<https://www.cdc.gov/flu/treatment/whatyoushould.htm>] (consulté le 21 février 2022)
- Commission Nationale de l'Informatique et des Libertés. Délibération n°2017-305 du 7 décembre 2017 (2017).
- DAWOOD, F. S., JAIN, S., FINELLI, L., SHAW, M. W., LINDSTROM, S., GARTEN, R. J., GUBAREVA, L. V., XU, X., BRIDGES, C. B. & UYEKI, T. M. 2009. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med*, 360, 2605-15.
- DAWOOD, F. S., IULIANO, A. D., REED, C., MELTZER, M. I., SHAY, D. K., CHENG, P. Y., BANDARANAYAKE, D., BREIMAN, R. F., BROOKS, W. A., BUCHY, P., FEIKIN, D. R., FOWLER, K. B., GORDON, A., HIEN, N. T., HORBY, P., HUANG, Q. S., KATZ, M. A., KRISHNAN, A., LAL, R., MONTGOMERY, J. M.,

- MØLBAK, K., PEBODY, R., PRESANIS, A. M., RAZURI, H., STEENS, A., TINOCO, Y. O., WALLINGA, J., YU, H., VONG, S., BRESEE, J. & WIDDOWSON, M. A. 2012. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. *Lancet Infect Dis*, 12, 687-95.
- EL ALLAKI, F., CHRISTENSEN, J. & VALLIÈRES, A. 2015. Comparing capture–recapture methods for estimation of the size of small and medium-sized populations using empirical data on commercial turkey farms in Canada. *Preventive Veterinary Medicine*, 120, 86-95.
- Équipes de surveillance de la grippe. 2022. Surveillance de la grippe en France, saison 2021-2022. *Bull Epidemiol Hebd*, 362-375.
- FARCOMENI, A. & DOTTO, F. 2021. A correction to make Chao estimator conservative when the number of sampling occasions is finite. *Statistics & Probability Letters*, 176, 109154.
- GALLAY, A., NARDONE, A., VAILLANT, V. & DESENCLOS, J. C. 2002. La méthode de capture-recapture appliquée à l'épidémiologie : principes, limites et applications. *Rev Epidemiol Sante Publique*, 50, 219-32.
- GIRAUD, T. 2017. *mapsf: Thematic Cartography*. R package version 0.6.1. [en ligne] [<https://CRAN.R-project.org/package=mapsf>]
- GOHEL, D., SKINTZOS, P., (2022). *flextable: Functions for Tabular Reporting*. R package version 0.8.3. [en ligne] [<https://CRAN.R-project.org/package=flextable>]
- GOUDIE, I. B. J. & GOUDIE, M. 2007. Who captures the marks for the Petersen estimator? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170, 825-839.
- GrippeNet. *Présentation du projet* [en ligne]. [<https://www.grippenet.fr/fr/covidnet/>] (consulté le 4 janvier 2023).
- GROTHENDIECK, G. 2017. *sqldf: Manipulate R Data Frames Using SQL*. R package version 0.4-11. [en ligne] [<https://CRAN.R-project.org/package=sqldf>]
- HASSEL, M., ASBJØRNSLETT, B. E. & HOLE, L. P. 2011. Underreporting of maritime accidents to vessel accident databases. *Accident Analysis & Prevention*, 43, 2053-2063.
- HUBERT, B. & DESENCLOS, J. C. 1993. Evaluation de l'exhaustivité et de la représentativité d'un système de surveillance par la méthode de capture-recapture. Application à la surveillance des infections à méningocoque en France en 1989 et 1990. *Rev Epidemiol Sante Publique*, 41, 241-9.
- Institut National de la Statistique et des études économiques (INSEE). 2023. *Estimation de la population au 1<sup>er</sup> janvier 2023. Séries par région, département, sexe et âge de 1975 à 2023*. [en ligne] [<https://www.insee.fr/fr/statistiques/1893198>]
- Institut Pasteur. *Missions du CNR des virus des infections respiratoires (dont la grippe)* [en ligne]. Mis à jour le 17/10/2018. [<https://www.pasteur.fr/fr/sante-publique/centres-nationaux-referance/cnr/virus-infections-respiratoires-dont-grippe/missions-du-cnr-virus-infections-respiratoires-dont-grippe>] (consulté le 3 janvier 2023)
- International Working group for Disease Monitoring and Forecasting (IWGDMF). 1995. Capture-recapture and multiple-record systems estimation I: History and theoretical development. *Am J Epidemiol*, 142, 1047-58.
- JOHNSON, N. P. & MUELLER, J. 2002. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bulletin of the history of medicine*, 76, 105-115.
- KRAMMER, F., SMITH, G. J. D., FOUCHIER, R. A. M., PEIRIS, M., KEDZIERSKA, K., DOHERTY, P. C., PALESE, P., SHAW, M. L., TREANOR, J., WEBSTER, R. G. & GARCÍA-SASTRE, A. 2018. Influenza. *Nat Rev Dis Primers*, 4, 3.
- LA RUCHE, G., DEJOUR SALAMANCA, D., BERNILLON, P., LEPARC GOFFART, I., LEDRANS, M., ARMENGAUD, A., DEBRUYNE, M., DENOYEL, G. A., BRICHLER, S., NINOVE, L., DESPRES, P. & GASTELLU ETCHEGORRY, M. 2013. Capture-recapture method for estimating annual incidence of imported dengue, france, 2007-2010. *Emerg Infect Dis*, 19, 1740-8.
- LOURY, P., JONES, G., CHAPPERT, J. L., PIVETTE, M., HUBERT, B. & GROUPE, E. 2017. Analyse de l'exhaustivité et de la qualité de la surveillance des gripes sévères, 2009-2013. Saint-Maurice: Santé publique France.

- MCNAIR, J. N., RUETZ, C. R., 3RD, CARLSON, A. & SUH, J. 2018. Reducing effects of dispersal on the bias of 2-sample mark-recapture estimators of stream fish abundance. *PLoS One*, 13, e0200733.
- Ministère de la Santé et de la Prévention. *Financement des établissements de santé* [en ligne]. Mise à jour le 23 février 2022. [<https://sante.gouv.fr/professionnels/gerer-un-etablissement-de-sante-medico-social/financement/financement-des-etablissements-de-sante-10795/article/financement-des-etablissements-de-sante>] (consulté le 22 février 2023)
- Organisation Mondiale de la Santé (OMS). *Influenza (seasonal)* [en ligne]. Mis à jour le 31 mars 2018. [[https://www.who.int/fr/news-room/fact-sheets/detail/influenza-\(seasonal\)](https://www.who.int/fr/news-room/fact-sheets/detail/influenza-(seasonal))] (consulté le 3 janvier 2023)
- PAPOZ, L., BALKAU, B. & LELLOUCH, J. 1996. Case counting in epidemiology: limitations of methods based on multiple data sources. *Int J Epidemiol*, 25, 474-8.
- PIVETTE, M., LOURY, P. & ET LES ÉPIDÉMIOLOGISTES EN CHARGE DE LA SURVEILLANCE DES GRIPPES, S. 2019. Analyse de l'exhaustivité de la surveillance des gripes sévères en France métropolitaine, saison 2017-2018. *Bull Epidemiol Hebd*, 571-572.
- Protocole Ecsir 2018. Version du 24 mai 2018. (Document interne)
- Réseau Sentinelles. Bilan d'activité 2021. [en ligne] [[5740 \(sentiweb.fr\)](https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/articles/la-grippe-une-epidemie-saisonniere)]
- RIVEST, L.-P. & BAILLARGEON, S. 2007. Applications and Extensions of Chao's Moment Estimator for the Size of a Closed Population. *Biometrics*, 63, 999-1006.
- Santé publique France, a. *La grippe, une épidémie saisonnière* [en ligne]. Publié le 22 avril 2019. [<https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/articles/la-grippe-une-epidemie-saisonniere>] (consulté le 3 janvier 2023)
- Santé publique France, b. *La surveillance épidémiologique de la grippe* [en ligne]. Publié le 22 avril 2019. [<https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/articles/la-surveillance-epidemiologique-de-la-grippe>]
- Santé publique France, c. Protocole de la surveillance sentinelle des cas graves de grippe et de COVID-19 nécessitant une prise en charge en réanimation, année 2020-2021. Version 2. 25 mars 2021. [en ligne] [<https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/articles/le-dispositif-de-surveillance-de-la-grippe>] (consulté le 21 février 2023)
- Santé publique France, d. *Réseau OSCOUR®* [en ligne]. Mis à jour le 25 novembre 2021. [<https://www.santepubliquefrance.fr/surveillance-syndromique-sursaud-R/reseau-oscour-R>] (consulté le 3 janvier 2023)
- Santé publique France, e. *Réseau SOS Médecins* [en ligne]. Mis à jour le 25 novembre 2021. [<https://www.santepubliquefrance.fr/surveillance-syndromique-sursaud-R/reseau-sos-medecins>] (consulté le 4 janvier 2023)
- Santé publique France, f. *Le dispositif de surveillance de la grippe* [en ligne]. Mis à jour le 18 novembre 2022. [<https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/articles/le-dispositif-de-surveillance-de-la-grippe>] (consulté le 4 janvier 2023)
- Santé publique France. Bulletin hebdomadaire Grippe-Semaine 01 (2023). [en ligne]
- SIEVERT, C. 2020. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC Florida
- SIEVERT, C., IANNONE, R., ALLAIRE, J.J., BORGES, B. (2023). *flexdashboard: R Markdown Format for Flexible Dashboards*. R package version 0.6.1. [en ligne] [<https://CRAN.R-project.org/package=flexdashboard>]
- SJOBERG DD, WHITING K, CURRY M, LAVERY JA, LARMARANGE J. 2021. Reproducible summary tables with the gtsummary package. *The R Journal* 2021 ;13:570–80.
- SPACCAFERRI, G., CAZEIN, F., LIEVRE, L., GEFFARD, S., GALLAY, A. & PILLONEL, J. 2010. Estimation de l'exhaustivité de la surveillance du sida par la méthode capture-recapture, France, 2004-2006. Saint-Maurice: Institut de veille sanitaire.
- TILLING, K. 2001. Capture-recapture methods—useful or misleading? *International Journal of Epidemiology*, 30, 12-14.

- TYERS, M. 2021. recapr: Two Event Mark-Recapture Experiment. R package version 0.4.4. [en ligne] [<https://CRAN.R-project.org/package=recapr>]
- UYEKI, T. M., HUI, D. S., ZAMBON, M., WENTWORTH, D. E. & MONTO, A. S. 2022. Influenza. *Lancet*, 400, 693-706.
- VAN HEST, N. A. H., GRANT, A. D., SMIT, F., STORY, A. & RICHARDUS, J. H. 2008. Estimating infectious diseases incidence: validity of capture–recapture analysis and truncated models for incomplete count data. *Epidemiology & Infection*, 136, 14-22.
- VIBOUD, C., ALONSO, W. J. & SIMONSEN, L. 2006. Influenza in tropical regions. *PLoS Med*, 3, e89.
- WEBSTER, R. G. & GOVORKOVA, E. A. 2014. Continuing challenges in influenza. *Ann N Y Acad Sci*, 1323, 115-39.
- WICKHAM, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York
- WITTES, J. & SIDEL, V. W. 1968. A generalization of the simple capture-recapture model with applications to epidemiological research. *J Chronic Dis*, 21, 287-301.
- WITTES, J. T., COLTON, T. & SIDEL, V. W. 1974. Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *J Chronic Dis*, 27, 25-36

## Annexes

### Annexe 1 – Fiche de signalement des cas sévères de grippe admis en réanimation dans le cadre de la surveillance SGS



**Surveillance des cas graves de GRIPPE et COVID-19 admis en réanimation**  
A remplir pour tout cas correspondant à la définition de cas et à adresser à la **Cellule régionale XXX**  
par messagerie de santé sécurisée: **MSSANTE: xxx@santepubliquefrance.mssante.fr**  
ou à déposer sur la **page de dépôt sécurisée:**  
<https://bluefiles.orange-business.com/xxx>

**Définition de cas :** Patient présentant une confirmation biologique de grippe ou d'infection à SARS-CoV-2 nécessitant une prise en charge en réanimation du fait de l'infection ou du fait de ses conséquences (décompensation de pathologie sous-jacente, surinfection bactérienne) – hors cas de syndrome inflammatoire multi-systémique pédiatrique lié à SARS-CoV-2 (Paediatric Inflammatory multisystem syndrome – PIMS)

Hôpital : ..... Service : ..... Tél. : .....

Médecin déclarant : ..... E-mail : .....

**Patient** Nom : ..... Prénom : ..... Date de naissance : \_\_\_/\_\_\_/\_\_\_ Sexe :  H  F

Pays de naissance : ..... Résidence → Dép.: ..... Commune : ..... Arr. : ..... (Lyon, Marseille, Paris)

Grossesse en cours  Professionnel de santé  Vit en établissement médico-social (ex : Ehpad, MAS, FAM, ...)

Fumeur actuel

Antécédent de Covid-19 confirmé (de plus de 2 mois) Date épisode précédent (DDS ou date du prélèvement) : \_\_\_/\_\_\_/\_\_\_

Vacciné contre la grippe pour la saison 2022-2023 (vaccination de l'entourage si nourrisson <6mois avec facteur(s) de risque de grippe sévère)

Vacciné contre la COVID-19 :  Oui  Non  ne sait pas Nombre de dose(s) reçue(s) : \_\_\_

Vaccin utilisé pour la dernière dose : ..... Date de la dernière dose reçue : \_\_\_/\_\_\_/\_\_\_

IMC : \_\_\_ kg.m<sup>2</sup>

Circonstance(s) d'exposition particulière(s) : ..... (ex : voyage, cluster identifié, ...)

Date de début des signes : \_\_\_/\_\_\_/\_\_\_ Date d'hospitalisation : \_\_\_/\_\_\_/\_\_\_ Date d'admission en réanimation : \_\_\_/\_\_\_/\_\_\_

Si le patient provient d'un autre service de réanimation, préciser l'hôpital et la commune : .....

Infection nosocomiale (grippe et/ou COVID-19)

Motif d'admission en réanimation :  Infection respiratoire aiguë virale  Décompensation de pathologie sous-jacente

Plusieurs réponses possibles

Surinfection bactérienne  Autre, préciser : .....

**Comorbidités (liste en annexe)**  Aucune  Ne sait pas

Pathologie cardiaque  Pathologie pulmonaire  Pathologie rénale  Pathologie hépatique

Pathologie neuromusculaire  Cancer (tumeur solide ou hémopathie) évolutif  Immunodépression  Diabète (type 1 et 2)

Hypertension artérielle (HTA)  Préciser : .....  Autre(s), préciser : .....

**Diagnostic virologique** Laboratoire ayant réalisé le diagnostic : ..... Date du prélèvement : \_\_\_/\_\_\_/\_\_\_

SARS-CoV-2 →  Positif  Négatif  Non recherché

Recherche de mutation/variants d'intérêt :  Non  Ne sait pas  Oui, par PCR  Oui, par séquençage

(criblage, Thermo Fisher, autre kit) (Sanger, NGS, WGS)

Si criblage :  E484K (A)  L452R (C)  « Mutation Omicron » (D) Autres : .....

Si séquençage :  Alpha  Beta  Gamma  Delta  Omicron  B.1.640 Autres : .....

Grippe →  Positif  Négatif  Non recherché

Si Positif :  A, sans précision  A(H1N1)  A(H3N2)  B, sans précision  B/Victoria  B/Yamagata  Non renseigné

Si coinfection virale confirmée, préciser le(s) virus identifié(s) : .....

#### A l'admission en réanimation - Pires valeurs lors de la première journée de séjour

Type d'assistance ou aide ventilatoire

Une seule réponse possible

O<sub>2</sub> (lunettes/masque)  VNI  O<sub>2</sub> haut débit  Ventilation invasive  Assistance extracorporelle (ECMO/ECCO2R)

Caractéristiques cliniques → FiO<sub>2</sub> : \_\_\_\_\_ % (ou débit d'O<sub>2</sub> : \_\_\_\_\_ L/min) PaO<sub>2</sub> : \_\_\_\_\_ mmHg (ou SpO<sub>2</sub> : \_\_\_\_\_ %)

**Evolution et bilan sur l'ensemble du séjour en réanimation – A renseigner à la sortie de réanimation**

Santé publique France – Cellule XXX  
Adresse postale – Standard +33 (0)xxx

[www.santepubliquefrance.fr](http://www.santepubliquefrance.fr)

Transfert vers un autre service de réanimation  
 Préciser l'hôpital et la commune si transfert : .....

Diagnostic d'embolie pulmonaire ou thrombose veineuse profonde :  Oui  Non

Sévérité maximale de SDRA<sup>2)</sup> au cours du séjour Une seule réponse possible : <sup>3)</sup> Adulte : selon critères de Berlin ; Enfant : selon définition PALICC Pediatr Crit Care Med 2015  
 Pas de SDRA  Mineur  Modéré  Sévère

Type(s) d'assistance ou aide ventilatoire : Plusieurs réponses possibles

Ventilation invasive

### Annexe: Liste détaillée des comorbidités

Type de pathologie	Définition ou exemple(s)
<b>Pathologie cardiaque</b>	Insuffisance cardiaque chronique Cardiopathies ischémiques Cardiopathie hypertensive Cardiopathie congénitale mal tolérée Valvulopathie et trouble du rythme grave
<b>Pathologie pulmonaire</b>	Insuffisance respiratoire chronique Bronchopneumopathie chronique obstructive (BPCO) Emphysème Fibrose du parenchyme pulmonaire Pneumoconiose Dysplasie broncho-pulmonaire Dilatation des bronches Mucoviscidose Asthme Syndrome d'apnée obstructive du sommeil (SAOS)
<b>Pathologie rénale</b>	Insuffisance rénale chronique Syndrome néphrotique
<b>Pathologie hépatique</b>	Cirrhose Atrésie biliaire Hépatite chronique
<b>Pathologie neuromusculaire</b>	Accident vasculaire cérébral (AVC) avec déficit séquellaire Pathologie neuromusculaire avec atteinte de la fonction respiratoire Pathologie neuro-développemental sévère (enfants)
<b>Pathologie cancéreuse</b>	Tumeur solide ou hémopathie maligne primitive en cours de traitement ou suivi médicalement

<b>Immunodépression</b>	Déficit immunitaire primitif ou acquis Asplénie anatomique ou fonctionnelle Hémoglobinopathie (drépanocytose, thalassémie majeure) Greffe d'organe Prise au long cours de médicaments immunosuppresseurs ou corticothérapie systémique Infection par le VIH
-------------------------	--

*La liste ci-dessus n'est pas exhaustive et est présentée afin d'aider au remplissage de la fiche clinique de signalement.*

Annexe 2 – Distribution des effectifs selon plusieurs variables communes aux bases SGS et PMSI pour les analyses restreintes aux services sentinelles

<b>Variables</b>	<b>SGS<sup>1</sup></b>	<b>PMSI<sup>1</sup></b>	<b>p-valeur</b>
<b>Effectif total</b>	475	1265	
<b>Sexe</b>			0,7 <sup>2</sup>
Femme	215 (45,3%)	558 (44,1%)	
Homme	260 (54,7%)	707 (55,9%)	
<b>Âge (médiane, EIQ)</b>	57 (20,5 – 67)	56 (12,0 – 69)	0,6 <sup>3</sup>
<b>Classe d'âge</b>			0,008 <sup>2</sup>
0-1 an	41 (8,6%)	161 (12,7%)	
2-17 ans	67 (14,1%)	205 (16,2%)	
18-64 ans	206 (43,4%)	449 (35,5%)	
65 ans et plus	161 (33,9%)	450 (35,6%)	
<b>Région</b>			<0,001 <sup>4</sup>
Auvergne-Rhône-Alpes	46 (9,7%)	184 (14,5%)	
Bourgogne-Franche-Comté	26 (5,5%)	47 (3,7%)	
Bretagne	32 (6,7%)	59 (4,7%)	
Centre-Val de Loire	26 (5,5%)	40 (3,2%)	
Corse	1 (0,2%)	9 (0,7%)	
Grand-Est	15 (3,2%)	67 (5,3%)	
Guadeloupe	3 (0,6%)	10 (0,8%)	
Guyane	3 (0,6%)	5 (0,4%)	
Hauts-de-France	50 (10,5%)	114 (9,0%)	
Île-de-France	54 (11,4%)	219 (17,3%)	
La Réunion	7 (1,5%)	15 (1,2%)	
Martinique	0 (0,0%)	4 (0,3%)	
Mayotte	13 (2,7%)	12 (0,9%)	
Normandie	26 (5,5%)	67 (5,3%)	
Nouvelle-Aquitaine	41 (8,6%)	90 (7,1%)	
Occitanie	39 (8,2%)	107 (8,5%)	
Pays de la Loire	69 (14,5%)	107 (8,5%)	
Provence-Alpes-Côte d'Azur	24 (5,1%)	109 (8,6%)	
<b>Type d'hôpital</b>			<0,001 <sup>2</sup>
CHU	276 (58,1%)	848 (67,0%)	
CH	194 (40,8%)	395 (31,2%)	
Autre	5 (1,1%)	22 (1,8%)	
<b>Unité médicale</b>			0,002 <sup>2</sup>
Réanimation adulte	377 (79,4%)	911 (72,0%)	
Réanimation pédiatrique	98 (20,6%)	354 (28,0%)	

<sup>1</sup>n (%); Médiane (EI)

<sup>2</sup>test du khi-deux d'indépendance

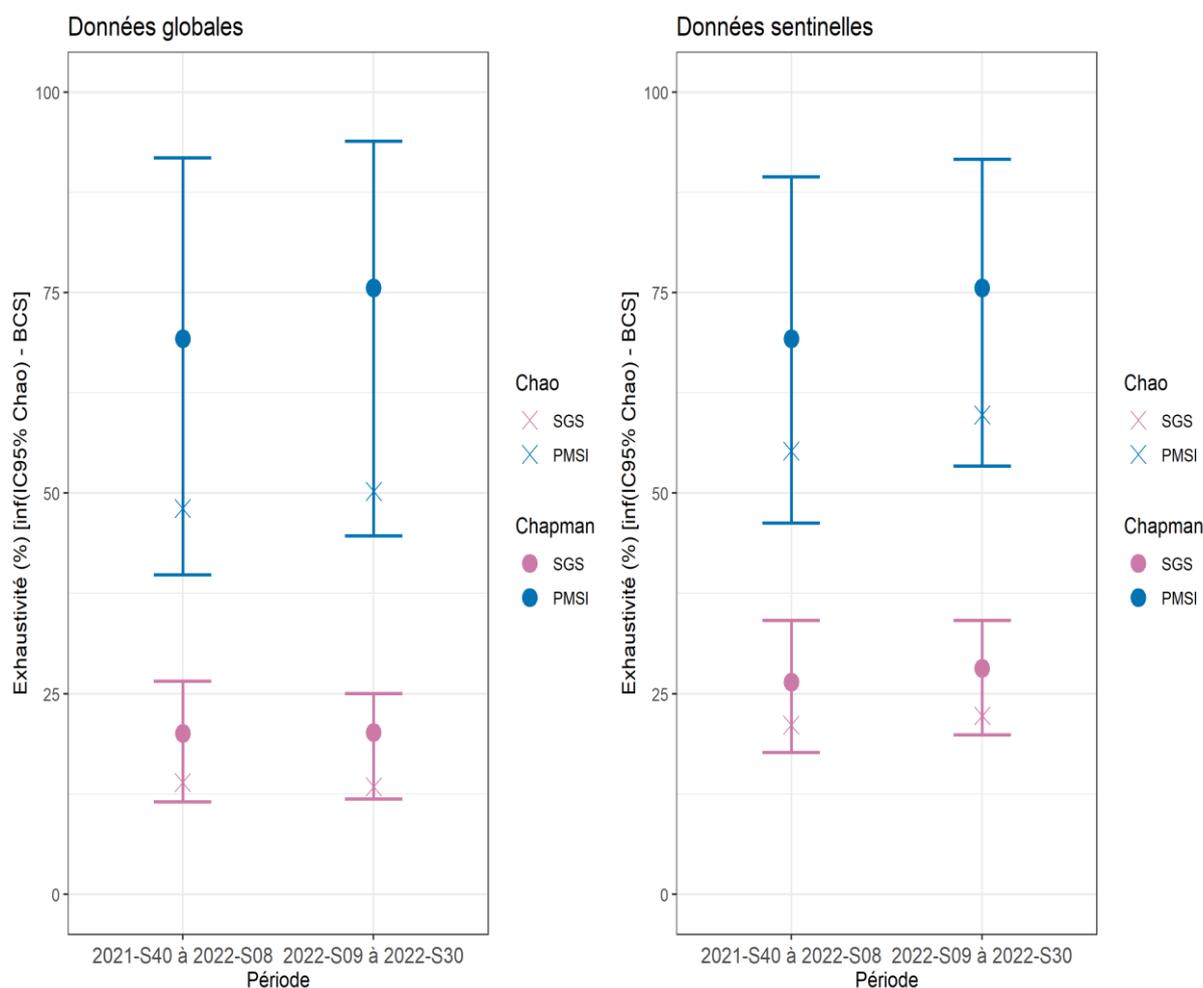
<sup>3</sup>test de Wilcoxon-Mann-Whitney

<sup>4</sup>test exact de Fisher avec p-valeur simulée (2000 itérations)

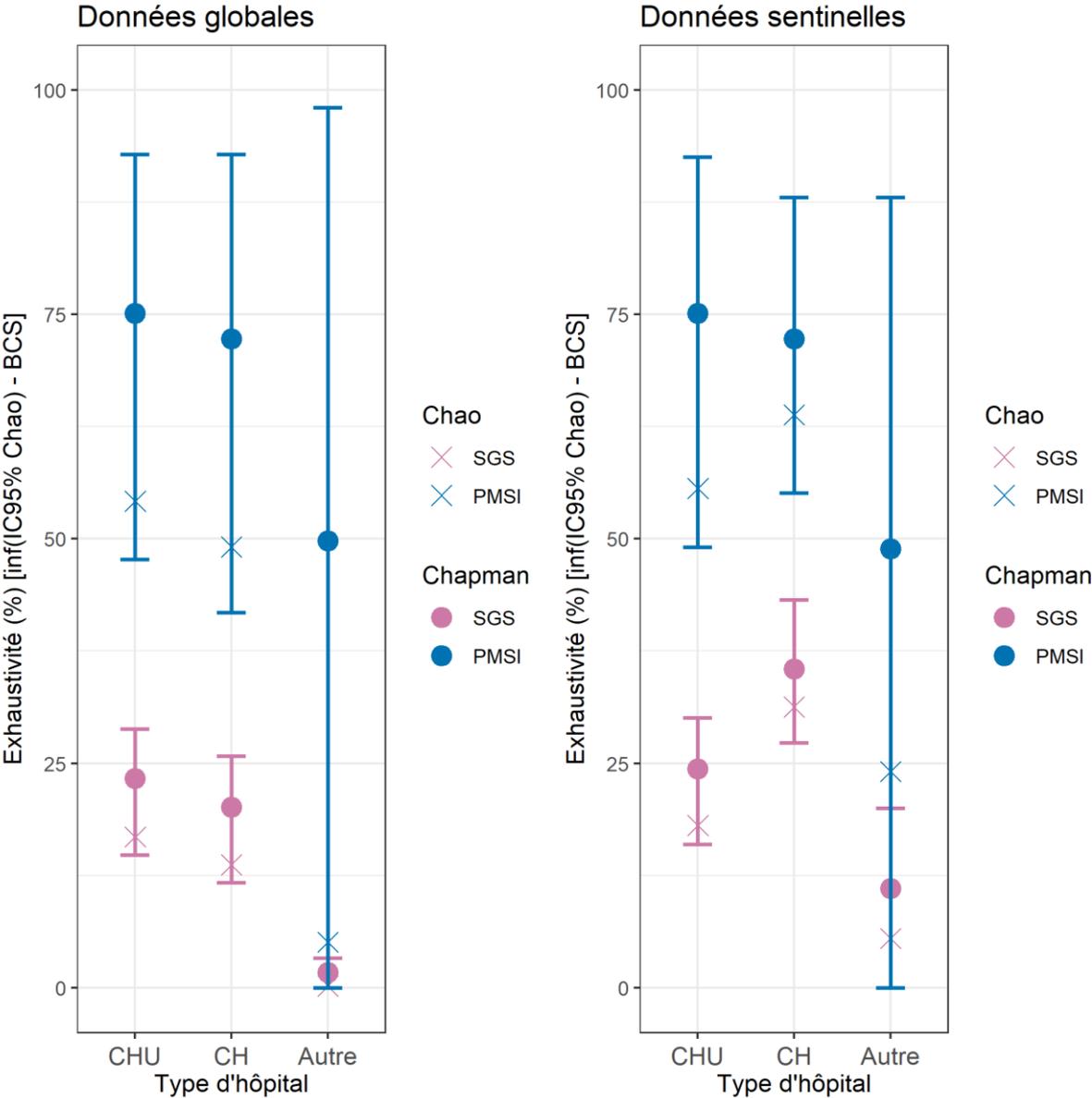
### Annexe 3 – Nombre de cas récupérés par chaque requête de croisement

Requêtes	Nombre de cas
<b>R1</b> <b>(égalité stricte)</b>	283
<b>R2</b> <b>(égalité stricte sans le département de domicile)</b>	21
<b>R3</b> <b>(différence d'âge de plus ou moins 1 autorisée)</b>	2
<b>R4</b> <b>(différence de date d'admission comprise entre moins 5 jours et plus 4 jours autorisée)</b>	42
<b>R5</b> <b>(différence d'âge de plus ou moins 1 et différence de date d'admission comprise entre moins 5 jours et plus 4 jours autorisées)</b>	1
<b>Total</b>	349

### Annexe 4 – Estimation de l'exhaustivité par source et par période de temps pour les analyses globales et restreintes aux services sentinelles, France entière, Semaine 40 de 2021 à Semaine 30 de 2022

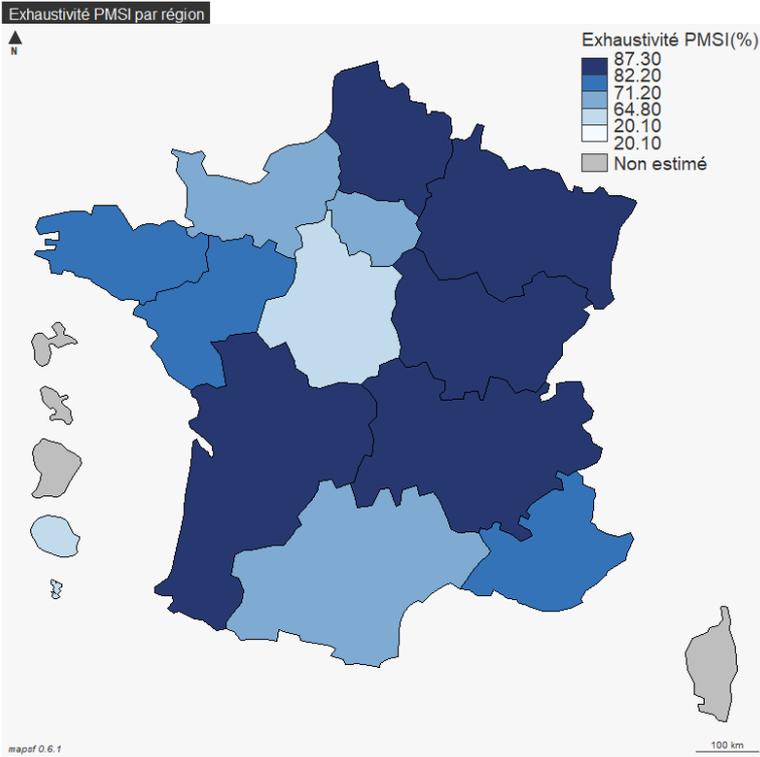
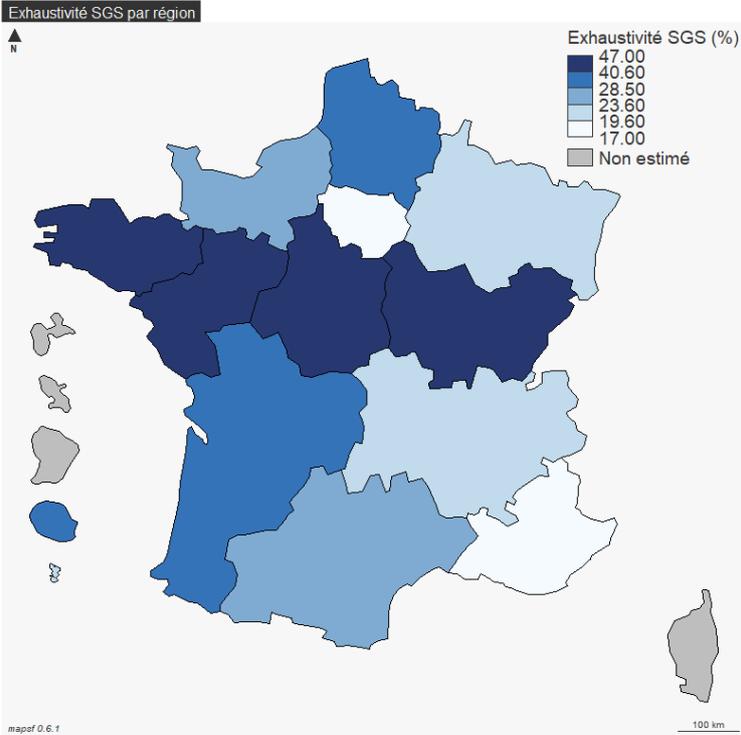


Annexe 5 – Estimation de l'exhaustivité par source et par type d'hôpital pour les analyses globales et restreintes aux services sentinelles, France entière, Semaine 40 de 2021 à Semaine 30 de 2022

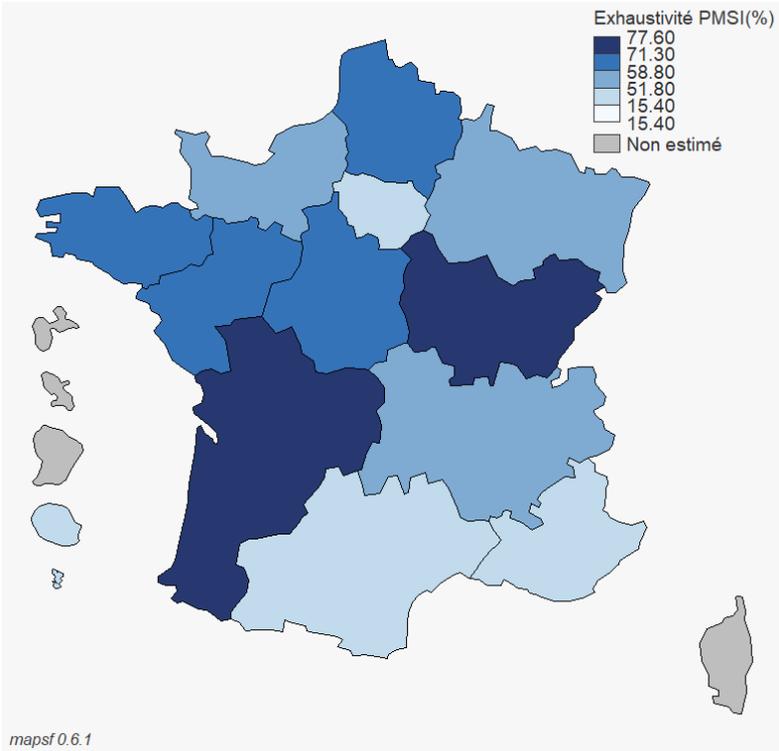
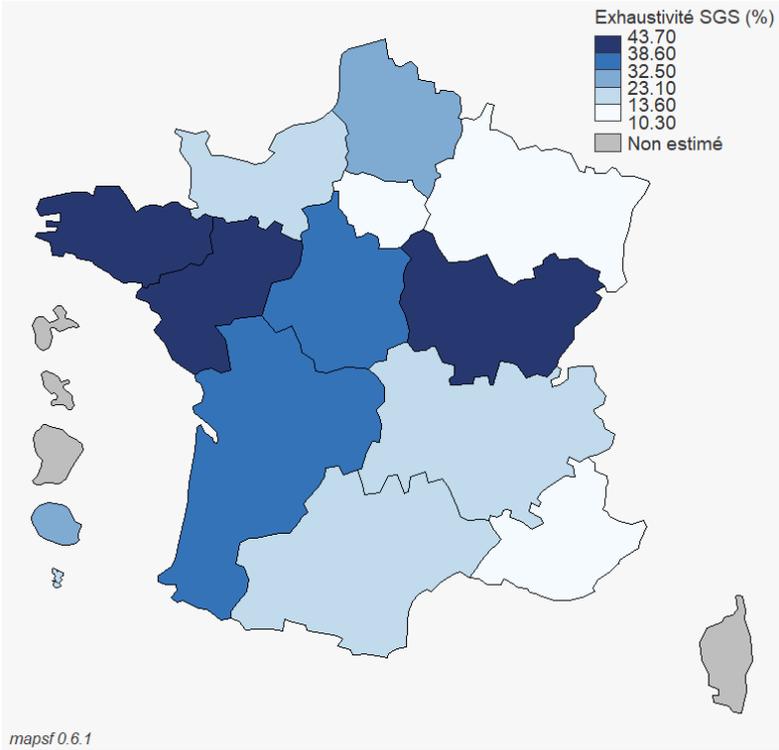


Annexe 6 – Cartes des exhaustivités moyennes SGS et PMSI pour les analyses restreintes aux services sentinelles

A - Exhaustivités estimées avec l'estimateur de Chapman



B – Exhaustivités estimées avec l'estimateur de Chao



## Annexe 7 – Estimation du nombre total de cas de grippe sévères admis en réanimation par stratification selon les estimateurs de Chapman et Chao

### A - Analyses globales

<b>Variables</b>	<b>Chapman</b>	<b>Chao</b>
<b>N estimé</b>	2361	3501
<b>IC 95% de N</b>	2157-2599	3199-3843
<b>Total stratifié par âge</b>	2348	3506
<b>Total stratifié par région*</b>	2306	4048
<b>Total stratifié par période</b>	2360	3502
<b>Total stratifié par type hôpital</b>	2468	7284

\*L'estimation à partir de l'analyse stratifiée par région ne prend pas en compte les régions où l'exhaustivité n'a pas été estimée (voir II.2.5).

### B – Analyses restreintes aux services sentinelles

<b>Variables</b>	<b>Chapman</b>	<b>Chao</b>
<b>N estimé</b>	1720	2168
<b>IC 95% de N</b>	1580-1888	1986-2380
<b>Total stratifié par âge</b>	1722	2216
<b>Total stratifié par région*</b>	1683	2306
<b>Total stratifié par période</b>	1721	2171
<b>Total stratifié par type hôpital</b>	1722	2240

\*L'estimation à partir de l'analyse stratifiée par région ne prend pas en compte les régions où l'exhaustivité n'a pas été estimée (voir II.2.5).

## Annexe 8 – Expression de l'estimateur de Chao corrigée pour les faibles effectifs

Expression générale (Rivest et Baillargeon, 2007) :

$$\hat{N} = S + \frac{(t - 1) * f_1(f_1 - 1)}{2t * (n_{11} + 1)}$$

Avec t le nombre de sources