

Correction des incidences non consolidées pour l'indicateur syndromes grippaux

Introduction

Le réseau Sentinelles estime chaque semaine l'incidence des syndromes grippaux à partir des données transmises par des médecins généralistes participants. Le système a été conçu de façon flexible pour les médecins : il n'impose pas de rythme pour l'envoi des données. L'envoi peut être réalisé quotidiennement, hebdomadairement et ce, n'importe quel jour de la semaine. Les médecins doivent uniquement respecter une période d'antériorité maximale de 12 jours pour déclarer les cas observés dans leur patientèle.

Ainsi, une déclaration réalisée par un médecin peut couvrir au maximum 12 jours. Les estimations d'incidence étant réalisées de manière hebdomadaire le lundi, selon le jour où le médecin déclare les cas, un maximum de trois semaines de consolidation peut être observé.

Ceci implique que l'incidence estimée le lundi (semaine $s+1$) pour la semaine précédente (semaine s) n'est potentiellement pas réalisée à partir de l'ensemble des données, certaines pouvant être collectées au cours de la semaine s ou de la semaine suivante (semaine $s+1$). La première estimation réalisée à $s+1$ est alors considérée comme non consolidée, c'est également le cas pour l'estimation réalisée à $s+2$. Il faut attendre 3 semaines pour obtenir une estimation consolidée, c'est-à-dire une estimation réalisée à l'aide de l'ensemble des données fournies par les médecins Sentinelles. On parle alors de consolidation.

Lors de l'analyse du processus de consolidation de l'estimation de l'incidence des syndromes grippaux, on observe que les estimations non consolidées (estimées à $s+1$), surestiment la valeur consolidée de l'incidence, (estimée à $s+3$), notamment lors des périodes épidémiques.

L'objectif est de corriger l'estimation non-consolidée de l'incidence durant les périodes épidémiques, afin d'obtenir une incidence « corrigée » la plus proche possible de l'incidence consolidée, estimée trois semaines plus tard.

Correction par régression multiple

Pour étudier ce phénomène de consolidation pour chaque semaine s , un indicateur appelé taux de variation a été utilisé :

$$\text{Taux de Variation } (s) = \frac{\text{Incidence}_{s+3}(s) - \text{Incidence}_{s+1}(s)}{\text{Incidence}_{s+1}(s)}$$

Où s est la semaine de l'estimation, $Incidence_{s+3}(s)$ est l'incidence pour la semaine s , estimée trois semaines plus tard (incidence consolidée) et $Incidence_{s+1}(s)$ est l'incidence pour la semaine s , estimée une semaine plus tard (incidence non-consolidée).

En période épidémique, l'évolution de ce taux de variation par semaine est semblable d'une année à l'autre. Une régression multiple est utilisée pour modéliser ce taux de variation, les variables explicatives étant le « logarithme de l'index de la semaine épidémique » et une variable binaire « vacances ».

L'équation de la régression modélisant le taux de variation T pour une semaine épidémique s est de la forme :

$$T(s) = a_0 + a_1 \times \log(s) + a_2 \times I_{vacances}(s) + \varepsilon(s)$$

Où a_0 est une constante, a_1 et a_2 des termes linéaires et $\varepsilon(s)$ l'erreur résiduelle. Les paramètres sont estimés par la méthode des moindres carrés. Pour estimer les paramètres du modèle une semaine s donnée, les données des 5 années passées sont utilisées.

A partir des paramètres estimés du modèle, l'incidence corrigée est estimée en multipliant le taux de variation prédit par le modèle par l'incidence non consolidée, auxquels on ajoute l'incidence non consolidée :

$$Incidence\ corrigée(s) = T(s) \times Incidence_{s+1}(s) + Incidence_{s+1}(s)$$

Pour calculer l'intervalle de confiance à 95% de cette incidence corrigée, il est nécessaire d'estimer sa variance. On remarque que l'incidence corrigée est le produit de deux variables aléatoires (taux de variation $T(s)$ et incidence non consolidée $Incidence_{s+1}(s)$).

Ainsi, la variance, pour la semaine s , peut s'exprimer de la façon suivante :

$$\begin{aligned} Var(incidence\ corrigée(s)) \\ = Var(Incidence_{s+1}(s)) \times Var(T(s)) + Var(Incidence_{s+1}(s)) \times E(T(s))^2 \\ + Var(T(s)) \times E(1 + Incidence_{s+1}(s))^2 \end{aligned}$$

Où la variance de l'incidence non consolidée est calculée en supposant que le nombre de cas déclaré par médecin par semaine suit une Loi de Poisson (<http://www.sentiweb.fr/?page=methodes&txt=466>) et la variance du taux de variation est issu de la régression présentée ci-dessous.