

Prise en compte des jours fériés dans la construction des indicateurs épidémiologiques basés sur les tests biologiques

Note méthodologique – Santé publique France – 15 avril 2021

Cette note a pour objectif de décrire comment prendre en compte l'effet des jours fériés dans le calcul d'indicateurs de suivi de l'épidémie de Covid-19. Ces indicateurs sont construits à partir des données du système d'information du dépistage (SI-DEP).

Indicateurs épidémiologiques

Les indicateurs épidémiologiques principaux sont les suivants :

- Taux de dépistage (TD) sur 7 jours glissants pour 100 000 habitants.
- Taux de positivité (TP) sur 7 jours glissants (en %)
- Taux d'incidence (TI) sur 7 jours glissants pour 100 000 habitants

Pour une unité géographique donnée et un intervalle de temps donné I , on note $T(I)$ le nombre de personnes testées dans l'intervalle I , n'ayant pas d'antécédent de test positif au cours des 60 derniers jours et résidant dans cette unité géographique. Parmi ces personnes testées, on note $P(I)$ le nombre de personnes testées positives. On note N le nombre d'habitants résidant dans cette unité géographique.

Les taux sur 7 jours glissants publiés le jour j sont définis par :

$$TD(j) = \frac{T(I)}{N} \times 100000$$

$$TP(j) = \frac{P(I)}{T(I)} \times 100$$

$$TI(j) = \frac{P(I)}{N} \times 100000$$

où l'intervalle I est l'ensemble des jours (en date de prélèvement) indexés de $j-9$ à $j-3$.

Problématique

Le choix de produire quotidiennement des indicateurs sur une période de 7 jours consécutifs a reposé sur la nécessité d'inclure l'ensemble des jours d'une semaine dans le calcul des indicateurs, étant donné que la réalisation des tests est variable d'un jour à l'autre au cours d'une semaine avec un facteur saisonnier hebdomadaire. C'est ce que montre la Figure 1 qui décrit la distribution du nombre total de personnes testées depuis le début de la mise en place de SI-DEP par jour de la semaine. On observe que l'activité est assez semblable du lundi au vendredi (avec un nombre plus

important le lundi et vendredi), plus faible le samedi et beaucoup plus faible le dimanche. Le Tableau 1 représente la distribution du nombre de personnes testées par jour de la semaine (en date de prélèvement) pour l'ensemble de l'historique. Lorsqu'un jour férié correspond à une journée d'activité comprise entre le lundi et le vendredi (lundi et vendredi compris), les indicateurs de taux d'incidence et de taux de dépistage calculés sur 7 jours glissants peuvent être impactés assez fortement. En effet, l'activité d'un jour férié est proche de celle d'un dimanche. Par ailleurs, le nombre de tests non réalisés le jour férié (et donc de positifs) a tendance à être reporté le jour qui suit (voire même les jours suivants).

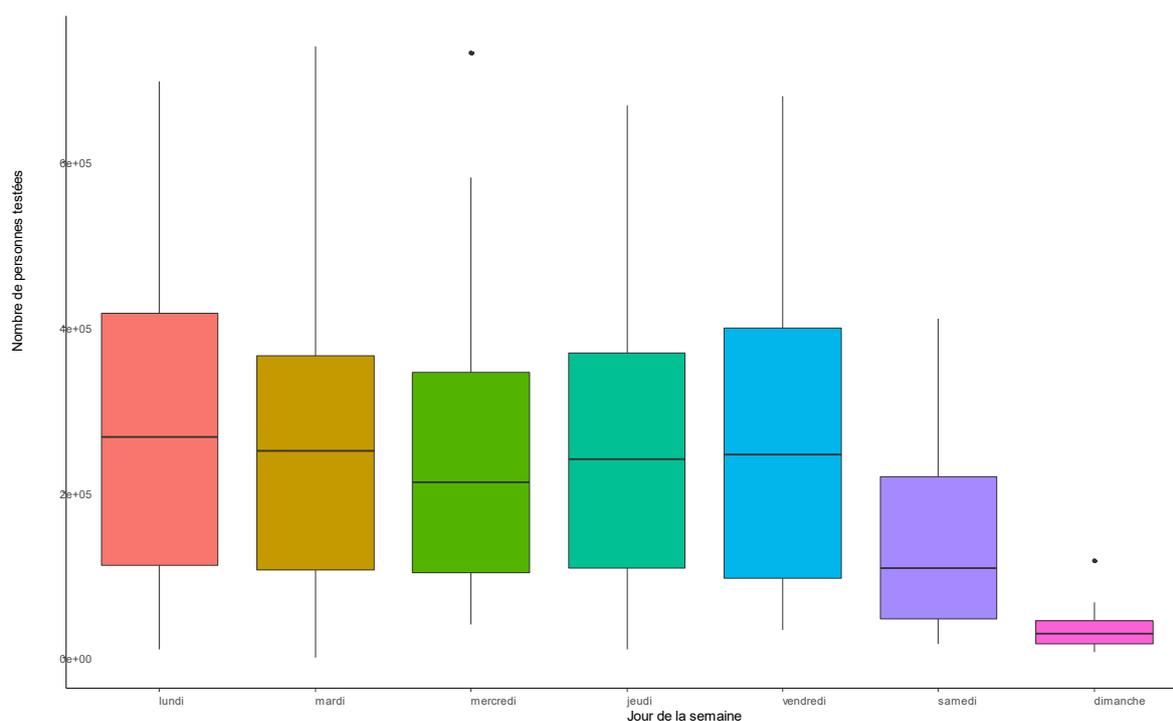


Figure 1 : Distribution du nombre de personnes testées par jour de la semaine (en date de prélèvement) depuis le début de la mise en place de SI-DEP.

Jour de la semaine (j)	Proportion (p(j))
Lundi	0.188
Mardi	0.177
Mercredi	0.164
Jeudi	0.175
Vendredi	0.179
Samedi	0.094
Dimanche	0.022
Total	1

Tableau 1 : Distribution du nombre de personnes testées par jour de la semaine (en date de prélèvement) pour l'ensemble de l'historique.

Ainsi, les nombres de personnes testées et de personnes testées positives au cours des 7 jours avec un jour férié sont mécaniquement plus faibles que pour une période de 7 jours ne comprenant pas de jour férié. Ces deux indicateurs

sont donc plus faibles et cette baisse, qui s'explique par une activité moindre, peut laisser penser à tort à un ralentissement voire à une baisse du nombre d'infections ou de dépistage.

Afin de réduire l'impact des jours fériés dans la valeur des indicateurs épidémiologiques et limiter les problèmes d'interprétation de leurs évolutions temporelles, nous avons appliqué une méthode de prise en compte des jours fériés.

Méthode proposée

L'impact des jours fériés dans le calcul des indicateurs épidémiologiques est connu dans le champ de la surveillance syndromique qui collecte et analyse quotidiennement des données transmises par des laboratoires ou des médecins. Une équipe anglaise de Public Health England a évalué cet impact et proposé une méthode de correction des jours fériés. Ce travail a été publié¹ dans *BMC Public Health* en 2017. Cette équipe a appliqué la méthode appelée *the extended working day moving average method* à différents regroupements syndromiques (asthme sévère, gastroentérites). La méthode est basée sur deux hypothèses : 1/ un jour férié est considéré comme étant équivalent à un dimanche et 2/ le premier jour ouvré après un jour férié est considéré comme un lundi. Le fait de considérer un jour férié comme un dimanche et le jour qui le suit comme un lundi consiste en pratique, à réaffecter les proportions présentées dans le Tableau 1.

Notons J, le jour de l'analyse. On note $TI(J) = \frac{P(I)}{N} \times 100000$, le taux d'incidence 7 jours glissants rapporté à 100 000 habitants, où I représente l'intervalle de temps [J-9 ;J-3] et N le nombre d'habitants pour la zone géographique considérée. On note p(i) la proportion de personnes testées le jour i dans une semaine : p(1) est la proportion de personnes testées le lundi, p(2) est la proportion de personnes testées le mardi, ..., p(7) est la proportion de personnes testées le dimanche. On a par définition, pour un intervalle de temps comprenant 7 jours consécutifs : p(1) +...+ p(7)=1.

On peut donc écrire le taux d'incidence 7 jours glissants comme : $TI(J) = \frac{P(I)}{p(1)+p(2)+p(3)+p(4)+p(5)+p(6)+p(7)} \times \frac{100000}{N}$.

L'hypothèse qu'un jour férié soit équivalent, en termes d'activités de déclaration, à un dimanche revient à remplacer la proportion associée au jour férié à la proportion associée au dimanche, c'est-à-dire p(7). Si par exemple le jour férié est un lundi, alors on remplace p(1) par p(7) dans la formule et le premier jour ouvré étant un mardi, on remplace p(2) par p(1). La formule du taux d'incidence devient alors : $TI(J) = \frac{P(I)}{p(7)+p(1)+p(3)+p(4)+p(5)+p(6)+p(7)} \times \frac{100000}{N}$. On observe donc que le dénominateur comporte 2 fois p(7), une proportion qui est assez faible. En reprenant les valeurs observées dans SI-DEP et décrites dans le Tableau 1, on obtient : p(7)+p(1)+p(3)+p(4)+p(5)+p(6)+p(7)=0.844. Le taux d'incidence corrigé du jour férié est donc 1/0.844=1,18 fois plus élevé que le taux d'incidence non corrigé du jour férié. Lorsque le jour férié est un mardi, un mercredi ou un jeudi, le même type de correction s'applique. En revanche, lorsqu'un vendredi est férié, on peut lui attribuer la proportion associée à un dimanche, mais on ne considèrera pas le samedi comme un lundi.

Une seule correction sur le vendredi sera appliquée. Si le samedi est férié on le considèrera comme un dimanche et si un dimanche est férié, aucune correction ne sera appliquée. D'autres situations peuvent survenir (2 jours fériés

¹ <https://bmcpublihealth.biomedcentral.com/articles/10.1186/s12889-017-4372-y>

consécutifs) et des attributions seront appliquées en fonction de ces situations, selon les jours de la semaine concernés.

Lorsque cette méthode est appliquée à des échelons géographiques plus fins (régions, départements), les proportions affectées à chaque jour de la semaine sont calculées pour chaque échelon géographique, afin de tenir compte de potentielles spécificités de déclaration pour chaque échelon.

La prise en compte des jours fériés a également un impact sur les taux de dépistage. En revanche le taux de positivité est inchangé puisque la même correction est appliquée au numérateur et au dénominateur.

Résultats

La Figure 2 illustre le taux d'incidence 7 jours glissants rapporté à 100 000 habitants en France en prenant en compte ou non les jours fériés. Les différences entre la courbe noire et la courbe rouge sont visibles lors de la présence de jours fériés (31/05/2020, 01/06/2020, 14/07/2020, 15/08/2020, 01/11/2020, 11/11/2020, 25/12/2020, 01/01/2021, 05/04/2021).

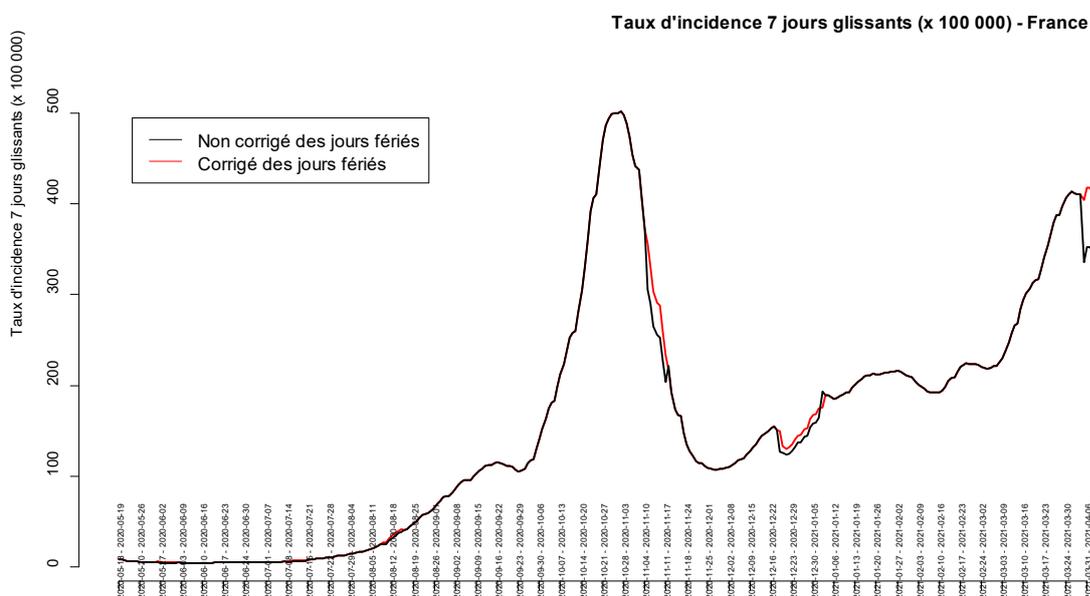


Figure 2 : Taux d'incidence 7 jours glissants en France, avec correction (courbe rouge) ou non (courbe noire) des jours fériés.

L'impact du lundi 5 avril 2021 est particulièrement marqué, d'une part parce qu'il s'agit d'un lundi qui est le jour de la semaine ayant l'activité de prélèvement la plus importante (Figure 1) et d'autre part parce que le taux d'incidence est élevé. A partir du 8 avril (jour J), la journée du 5 avril a été incluse dans l'intervalle de temps [J-9 ; J-3]. Le 5 avril, seulement 104559 personnes ont été testées en France contre en moyenne 560 000 tests au cours des 4 lundi précédents. Le taux d'incidence sur 7 glissants rapporté à 100 000 habitants est passé de 410 pour la période du 29/03 au 04/04 à 336 pour la période du 30/03 au 05/04, soit une baisse de 18%.

La Figure 3 illustre le taux de dépistage 7 jours glissants rapporté à 100 000 habitants en France en prenant en compte ou non les jours fériés.

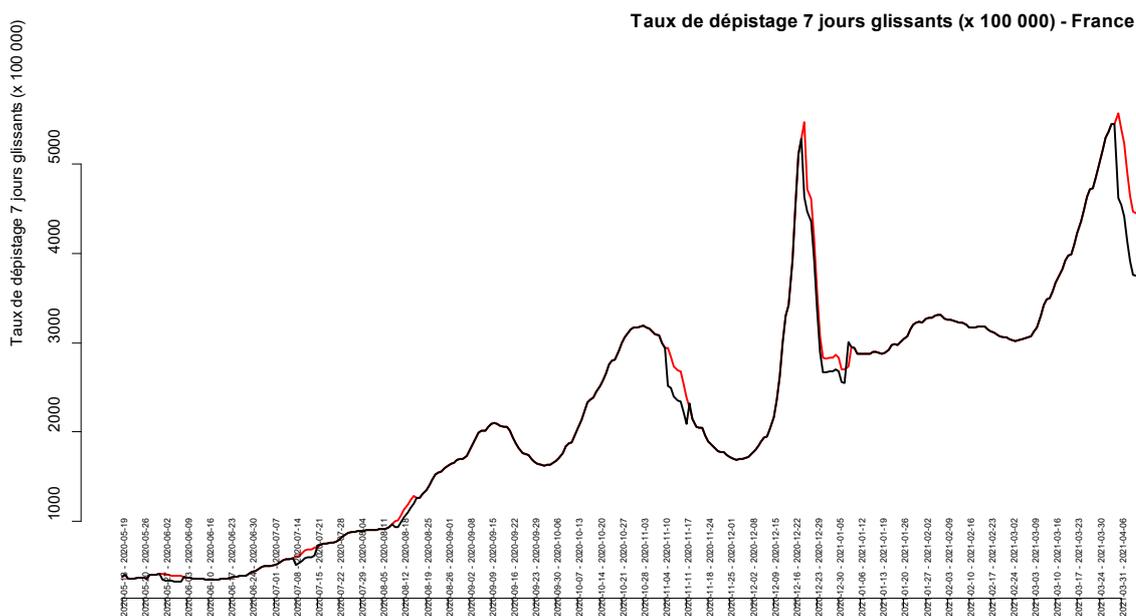


Figure 3 : Taux de dépistage 7 jours glissants en France, avec correction (courbe rouge) ou non (courbe noire) des jours fériés.

Concernant le taux de positivité (nombre de personnes testées positives divisé par le nombre de personnes testées), celui n'est pas affecté par les jours fériés car l'impact est le même pour le nombre de personnes positives (numérateur) et le nombre de personnes testées réalisés (dénominateur) et donc s'annulent ne nécessitant alors pas de corrections.

Discussion

La méthode appliquée est une méthode simple à implémenter, facile à expliquer et montre tout son intérêt pour le suivi de la dynamique de l'épidémie en cas de présence de jours fériés affectant l'activité de tests en France. Si elle est basée sur des hypothèses réalistes et probablement robuste (considérer un jour férié comme un dimanche), le choix de l'affectation du lundi, au jour suivant un jour férié, doit être questionné selon le jour de la semaine qui est férié. Lorsque plusieurs jours fériés se suivent des choix adaptés doivent également être pris au cas par cas. Il est également à noter qu'en cas de jour férié, des personnes peuvent anticiper la fermeture des laboratoires ou pharmacies en allant réaliser un test avant le jour férié. Dans ce cas de figure, l'effet jour férié est sans doute légèrement amoindri mais cette anticipation ne compense que marginalement l'effet du jour férié.

Enfin, il est important de noter que cette méthode permet d'estimer le nombre de personnes qui auraient été testées (et testées positives) s'il n'y avait pas eu de jour férié. Il s'agit donc d'une estimation d'un nombre de personnes testées et d'un nombre de cas confirmés biologiquement.

Les indicateurs épidémiologiques qui prennent en compte les jours fériés sont donc calculés à partir de tests biologiques validés auxquels sont ajoutés des tests « fictifs » sur la base des principes et hypothèses explicitées dans la section méthode.