

25 septembre 2020



Lettre n° 15

Données non-consolidées ; faux positifs ; caractère aléatoire des tests

Encore de petits pas vers la vérité...

Bonjour, ou bonsoir, si vous êtes à l'autre bout du monde.

Nous avons conclu la lettre précédente par une note optimiste basée sur l'évolution du taux d'incidence, défini comme le cumul sur 7 jours de nombre de testés positifs, par 100 000 habitants. Une dizaine de jours plus tard, les chiffres ont changé parce que les données de tests de Santé Publique France sont **révisées avec effet rétroactif**.

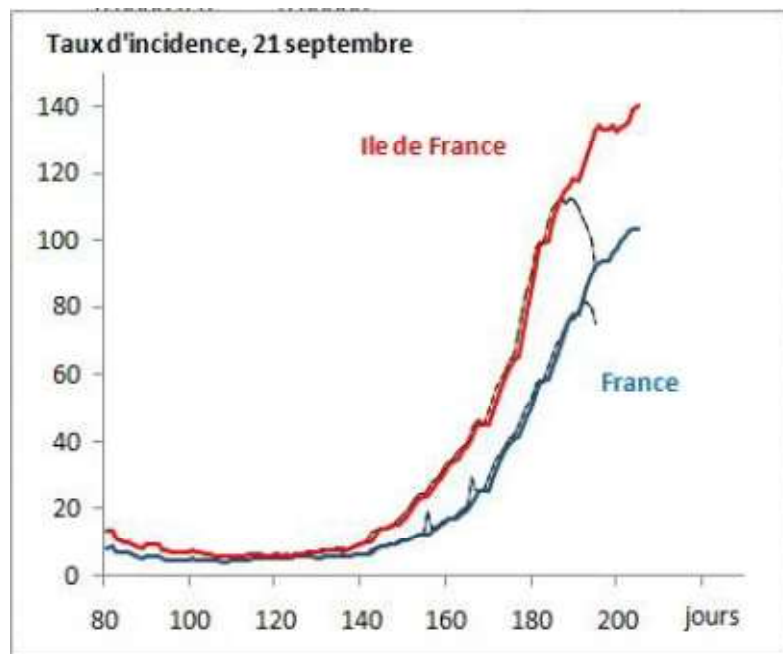
Nous avons trouvé l'avertissement suivant sur le site gouvernemental

<https://www.data.gouv.fr/fr/datasets/taux-d-incidence-de-lepidemie-de-covid-19/>

Précision :

- Le délai de remontée des tests peut excéder 9 jours dans certains cas. Les indicateurs sont ajustés quotidiennement selon la réception des résultats.

La comparaison, sur la figure ci-après, entre les anciennes (trait pointillé) et les nouvelles données (trait plein), est éloquent. La décroissance qui avait justifié notre optimisme n'est **pas confirmée**. Les données de tests étaient alors trop incomplètes, et correspondaient probablement, par malchance, à une partie moins représentative de la population testée.



Nous en tirons la leçon que l'indicateur « taux d'incidence » n'est pas fiable pour caractériser l'évolution récente. Il a de plus l'inconvénient d'être faussé par l'impact des tests sur la fraction visible de l'épidémie.

Le seul indicateur fiable et à jour de l'état de contamination de la population, ainsi que nous le répétons depuis la lettre n° 3, consisterait en campagnes répétées de dépistage sur des échantillons aléatoires. Ses données présenteraient le retard minimum par rapport à la contamination : temps d'analyse des tests + temps de remontée des données

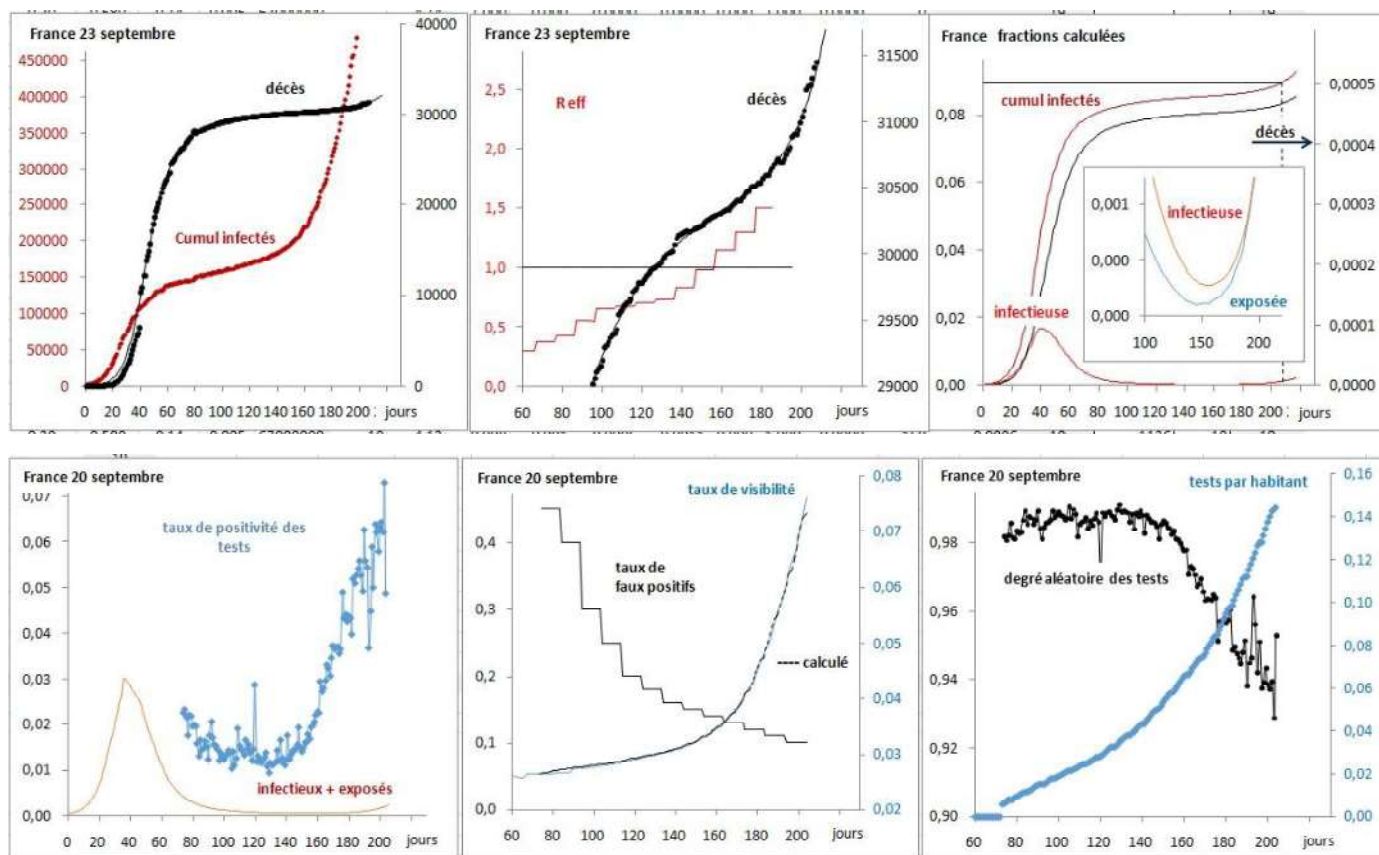
De même, le fléchissement de la croissance du taux d'incidence, déduite des données récentes, doit donc elle aussi être considérée avec circonspection.

Il faut donc attendre la consolidation de ces données. *On peut aussi se demander si la polémique actuelle concernant la Ville de Marseille n'est pas basée sur l'utilisation de telles données non-consolidées.*

Nous clarifions aussi une autre question restée pendante dans la précédente lettre, grâce à une remarque judicieuse de notre ami Henri L., qui a attiré notre attention sur les « **faux positifs** ». En effet, un test faussement positif ne conduit pas à la déclaration d'un cas, et donc ne contribue pas à augmenter le taux de visibilité de

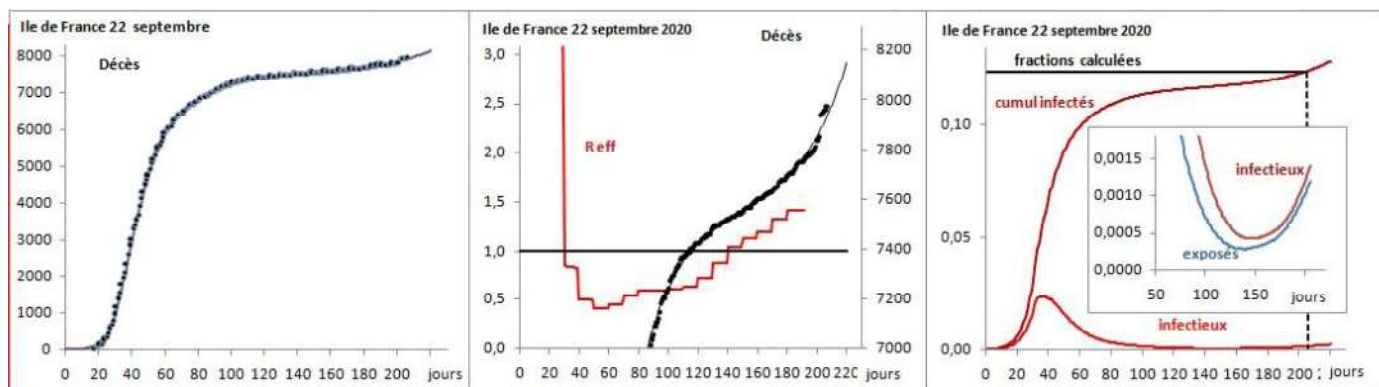
l'épidémie. Le facteur décrivant l'efficacité des tests à faire augmenter la visibilité est donc le taux de « vrais positifs » (égal à 1 - le taux de faux positifs). Il peut être évalué directement, au jour le jour, par le rapport nouveaux cas / nouveaux positifs. Nous avons choisi d'obtenir le taux de faux positifs au moyen d'un ajustement par paliers, représentant une moyenne sur plusieurs jours, des données de visibilité de l'épidémie. La visibilité calculée augmente chaque jour de la quantité : nouveaux tests positifs \times taux de vrais positifs divisée par la population infectée (égale à la population totale \times fraction IC calculée). Cette procédure avait déjà été décrite dans la lettre n° 10 (avec cependant une confusion de degré aléatoire des tests / taux de vrais positifs, déjà mentionnée).

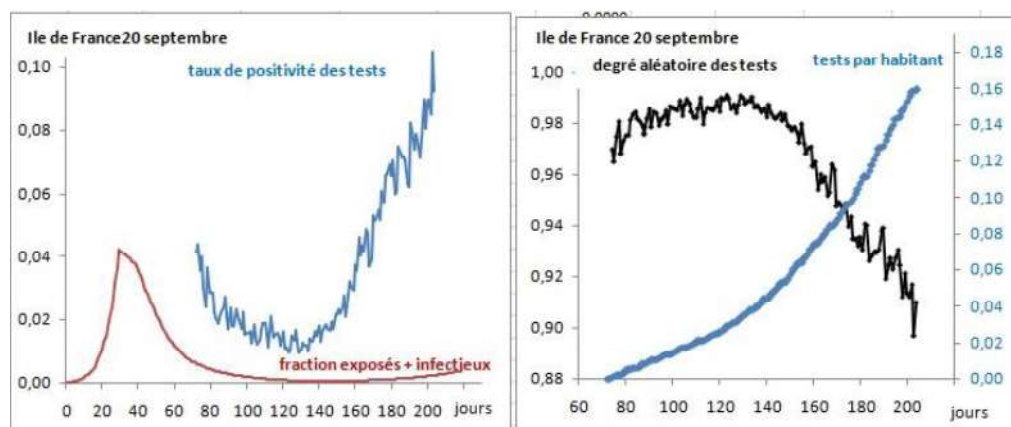
Reprenons les données de la France :



Pour l'échelle de temps : jour 1 = 1^{er} mars 2020 ; jour 62 = 1^{er} mai

Les tendances mentionnées dans la précédente lettre sont confirmées, sans signe notable d'amélioration de la situation. Et pour l'Ile de France, dont l'analyse est simplifiée par le fait que les données de cas déclarés ne sont plus disponibles depuis la fin Mars (pas d'éléments sur les faux positifs...) :



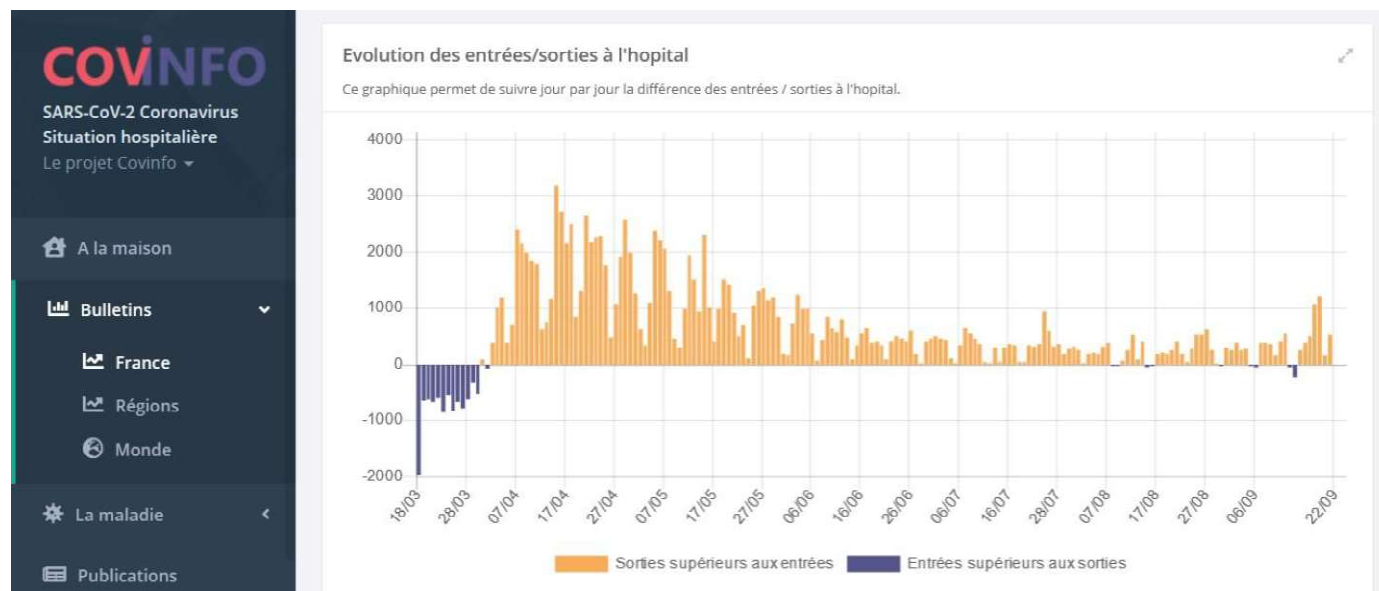


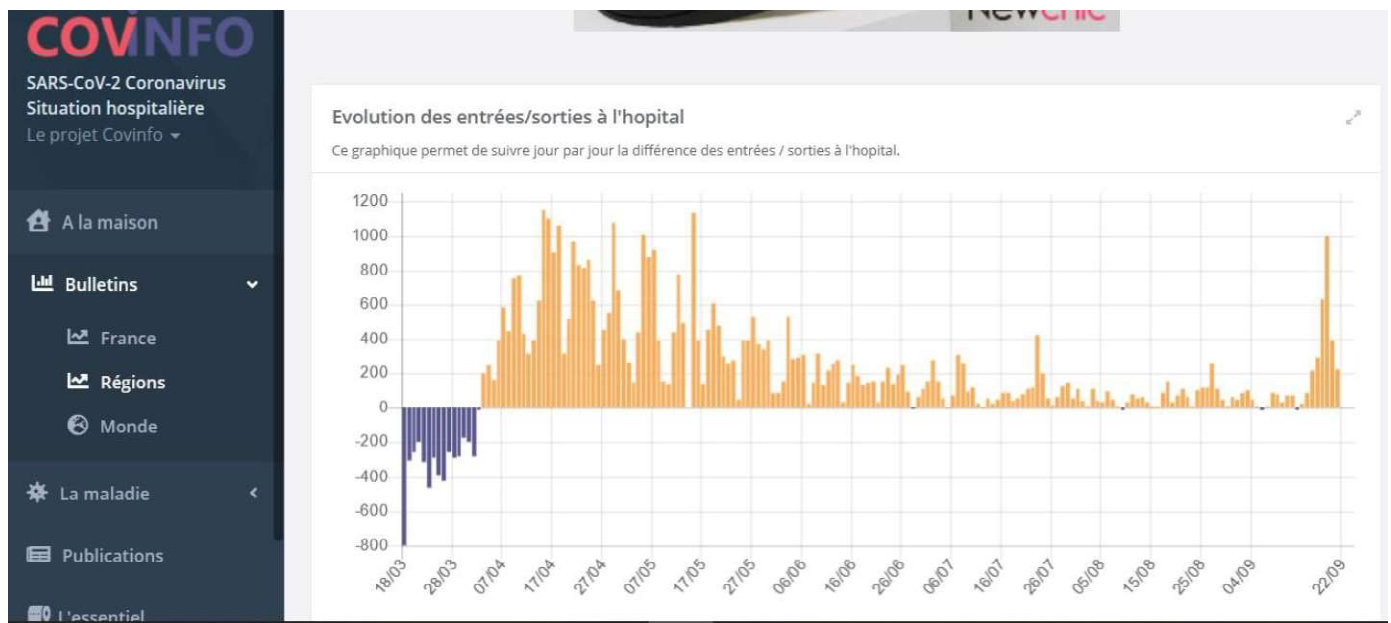
L'évolution du taux de positivité ne se ralentit pas. On restera réservé sur les données de tests des derniers jours, non-consolidées. Le caractère aléatoire des tests, nettement inférieur à la moyenne nationale, pourrait indiquer un ciblage encore plus affirmé de la campagne de dépistage.

L'accroissement subit du nombre de décès dans les derniers jours nous met face aux difficultés de notre analyse dynamique par le suivi de la courbe des décès. Nous étions déjà conscients du retard de cette information sur la dynamique instantanée de la propagation du virus. On rappellera que le suivi de l'épidémie par la courbe des cas (solution choisie par Santé Publique France), plus immédiat, présente l'inconvénient d'être affecté par le caractère massif de la campagne de tests. Nous assistons donc actuellement à un emballement de ce phénomène : davantage de tests => davantage de cas => incitation croissante à se faire tester.

Cet accroissement des décès annonce-t-il une aggravation durable de la situation, ou bien est-il une simple fluctuation sans lendemain, faisant suite à une accumulation temporaire de patients dans les services de réanimation ? On sait que les séjours y ont une durée très variable.

Pour tenter d'en savoir plus sur cette surmortalité nous avons consulté le site <https://www.covinfo.fr/region/ile-de-france> dont les données hospitalières sont tenues à jour et nous ont paru fiables (par contre les données de tests y sont arrêtées depuis que GEODES a pris le relais...). Voici pour la France :





La récente et brusque remontée des décès en Ile de France semble plutôt liée à un pic des entrées localisé sur les quelques jours précédents. Les personnes concernées pourraient donc avoir été atteintes plus gravement que la moyenne. L'avenir nous dira si cette « bulle » se reproduit ou non.

Portez-vous bien, en respectant les gestes barrière dans toute leur variété... (refrain connu).

François VARRET, Physicien, Professeur Émérite à l'Université de Versailles Saint-Quentin

Mathilde VARRET, Chargée de Recherche INSERM (Génétique, Biologie) Hôpital Bichat.