



HAL
open science

SARS-CoV-2 : Une simulation pré-confinement non-adéquate dans l'étude HAL Id: pasteur-02548181

Christian Manivel, Thierry Bourbié, Max Roche, Christian Chochillon,
Corinne Fernandez, Patrick Lagadec, Patrick Hontebeyrie

► To cite this version:

Christian Manivel, Thierry Bourbié, Max Roche, Christian Chochillon, Corinne Fernandez, et al.. SARS-CoV-2 : Une simulation pré-confinement non-adéquate dans l'étude HAL Id: pasteur-02548181. 2020. hal-02567189

HAL Id: hal-02567189

<https://hal.science/hal-02567189>

Preprint submitted on 7 May 2020

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

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

SARS-CoV-2: Une simulation pré-confinement non-adéquate dans l'étude HAL Id: pasteur-02548181

Auteur : Christian Manivel

Contributions : Thierry Bourbié, Max Roche, Christian Chochillon, Corinne Fernandez, Patrick Lagadec,
Patrick Hontebeyrie

Version v1

Ref : CML-SARS-CoV-2-01

Extrait de : Archive ouverte : HAL-Id : 02564410, v1 Preprint soumis le 5 mai 2020

Résumé

L'observation des données réelles des « décès par 24H à l'hôpital » de l'épidémie de SARS-CoV-2 en France nous a conduit à étudier quelles étaient les simulations faites pour les données pré-confinement dans différentes études publiées et en particulier dans l'étude identifiée HAL Id : pasteur-02548181, <https://hal-pasteur.archives-ouvertes.fr/pasteur-02548181>

L'observation des données réelles des « décès par 24H à l'hôpital » de l'épidémie de SARS-CoV-2 en France nous a conduit à constater que le confinement a eu un impact sur la mortalité à compter du 4 avril d'une part et d'en induire un décalage de 19 jours entre infection et mortalité issu de ces données à l'hôpital.

Nous avons donc pu confronter la simulation effectuée dans l'étude ci-dessus pour les données pré-confinement à la réalité des données, et nous avons conclu que cette simulation présentait une adéquation de 38 % avec la réalité.

« L'obsession de précision numérique sans attention aucune à la pertinence est le meilleur traceur de l'absence de culture mathématique véritable. » Carl Friedrich Gauss (1777 - 1855)

« A partir des données publiques, tout un chacun peut donc refaire les calculs ...Au lieu de raisonner sur des coefficients abstraits tel le fameux nombre moyen R_0 de contagions par personne, il faudrait pouvoir entrer dans le détail de ces contagions...» Hervé Le Bras, démographe – Le Monde – le 1^{er} mai 2020

SARS-CoV-2 : Une contagiosité qui décroît naturellement avec le temps

Christian Manivel

christian.manivel.1972@polytechnique.org

Sommaire

1 Observation des données réelles

1.1 Données brutes

1.2 Processus de collecte des données

1.3 Le confinement impacte la mortalité à compter du 4 avril

2 La simulation faite dans l'étude HAL Id 02548181

2.1 Les hypothèses de la simulation

2.2 La période de la simulation de 45 jours est cohérente

2.3 Les données réelles utilisées pour effectuer cette simulation

2.4 Le taux de mortalité

2.5 Les données de la simulation transcrites en mortalité

2.6 La comparaison Simulation vs. Réel

3 Conclusion

DOCUMENT PRINCIPAL

1 Observation des données réelles

1.1 Données brutes

Les données utilisées sont les données « décès à l'hôpital » issues de Santé Publique France.
<https://dashboard.covid19.data.gouv.fr>

Les données mortalité à l'hôpital ne sont pas la totalité des décès en France dus au virus, il manque la mortalité en institutions (EPHAD et EMS par exemple) et à la maison. Cependant la mortalité à l'hôpital représente une part importante de ces décès et on fera l'hypothèse que le comportement de cette population face au virus est homogène et que cette population est représentative de l'épidémie en cours.

Nous avons utilisé les données de mortalité par 24H pour faire cette étude d'adéquation entre les données réelles et la simulation effectuée dans l'étude HAL Id: pasteur-02548181 car les données de mortalité cumulées (intégration dans le temps de la courbe précédente) aurait naturellement un effet de lissage des données et fausserait en quelque sorte la comparaison.

Par ailleurs ces « données de mortalité par 24H à l'hôpital » sont le pendant, comme nous le verrons par la suite des « Daily new infections » de la simulation effectuée dans l'étude pré-citée pour les données pré-confinement.

1.2 Processus de collecte de données

L'observation de ces données brutes indique que le processus de collecte des données n'est pas stable.

En effet, il y a un déficit de collecte manifeste et chronique tous les samedi et dimanche de chaque semaine, avec un rattrapage le lundi et sans doute le mardi qui suit le week-end.

Un retraitement des données brutes pourrait être envisagé.
Il ne ferait qu'accroître la pertinence de ce qui est décrit ci-après.

Nous avons donc fait le choix de garder les données brutes, incontestables, comme base de notre étude.

1.3 Le confinement impacte la mortalité à partir du 4 avril

Une question immédiate se pose : à partir de quand les effets du confinement se sont-ils fait sentir sur la mortalité à l'hôpital ?

L'observation des données réelles nous permet d'affirmer sans grand risque que les données journalières de mortalité (nouveaux décès en 24H) ont cru sans interruption jusqu'au 3 avril, puis, en tendance, ont continué à baisser régulièrement.

Ce dernier fait n'est pas imputable au processus de collecte de données.

On peut donc prendre comme hypothèse crédible que les premiers effets du confinement sur la mortalité ont eu lieu le 4 avril 2020.

Dès lors le confinement ayant eu lieu le 16 mars, et ses premiers effets significatifs sur les infections le 17 mars, on peut en conclure (ce qui est cohérent avec ce que l'on sait à ce jour du processus incubation/infection/décès) que :

le décalage dans le temps entre infections et les décès corrélatifs est de 19 jours.

Ainsi, toutes les données réelles de mortalité jusqu'au 3 avril sont des données que l'on peut réputer correspondre à une épidémie en évolution libre, non affectées par le confinement.

Les données réelles de la mortalité par 24H à jour au 29 avril sont les suivantes : (Figure 1)

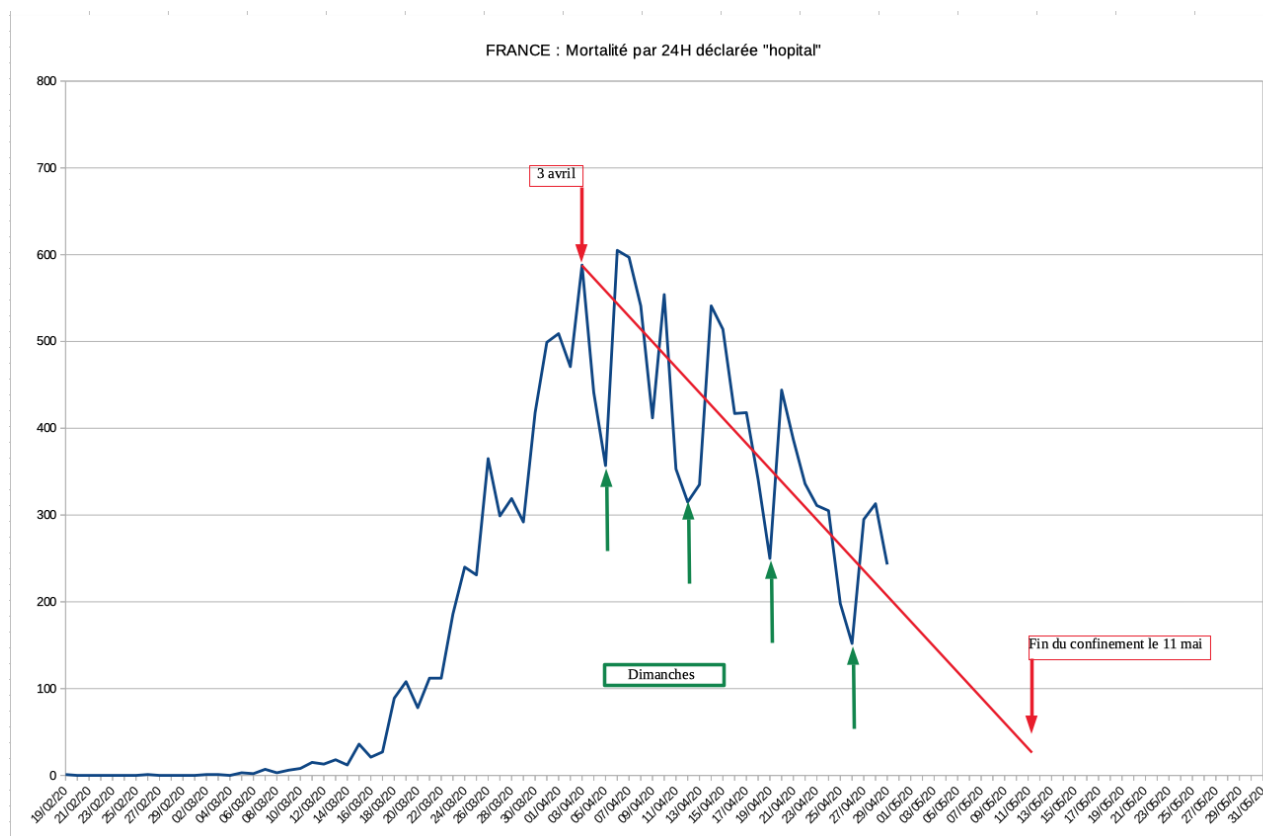


Figure 1

2 La simulation pré-confinement faite dans l'étude HAL Id 02548181

2.1 Les hypothèses de la simulation

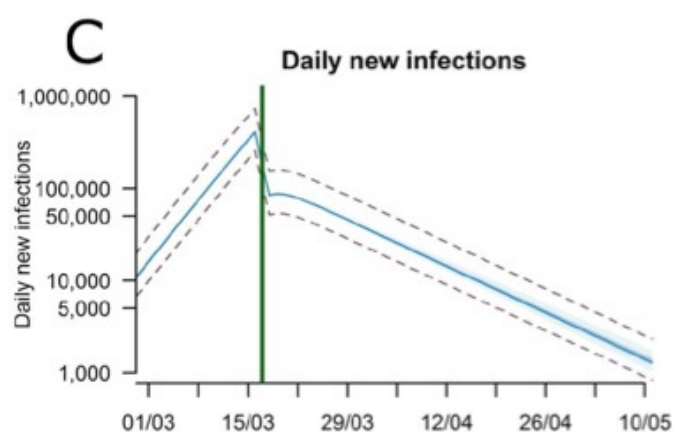
Les hypothèses de simulation de cette étude HAL Id 02548181 sont les suivantes :

« *Simulation study to assess model performance in estimating IFR and hospitalization risk ...For a period of 45 days we simulate a growing epidemic, seeded by a single infection, where the number of cases grows exponentially each day with an exponential growth rate of 0.3... » .*

Cette simulation a été effectuée par lissage sur la période du 1^{er} février au 17 mars.

La méthode de lissage utilisée consiste en une présentation des données pour les infections par 24H (Daily new infections) en échelle logarithmique base 10 et droite de régression comme meilleure approximation avec coefficient de croissance de 0,3.

Le résultat de cette simulation est présentée dans l'étude HAL Id 02548181:



2.2 La période de simulation de 45 jours est cohérente

La date de début de la période de simulation utilisée dans l'étude pour les nouvelles infections par 24H est le 1^{er} février (réputée correspondre à une valeur initiale de 1 cas à cette date).

Le premier décès identifié a eu lieu le 19 février.

Ceci est cohérent avec notre observation des données réelles et hypothèse que l'impact du confinement sur la mortalité avait eu lieu 19 jours après le début du confinement. (cf. paragraphe 1.3 ci-dessus).

2.3 Les données réelles utilisées pour créer cette simulation

Les données réelles utilisées dans cette étude sont bien les données publiques déjà citées ci-dessus.

2.4 Le taux de mortalité

Le taux de mortalité résultat de l'étude est de 0,53 %.

Il résulte de l'analyse des données réelles issue du Diamond Princess, pris comme échantillon représentatif de l'épidémie.

2.5 Les données de la simulation transcrites en mortalité

Dès lors, la simulation effectuée (Daily new infections) est transposable immédiatement en données du type Mortalité par 24H de la manière suivante :

- période de la simulation décalée de 19 jours :

La période pour les infections du 1^{er} février au 17 mars devient la période du 19 février au 3 avril pour la mortalité.

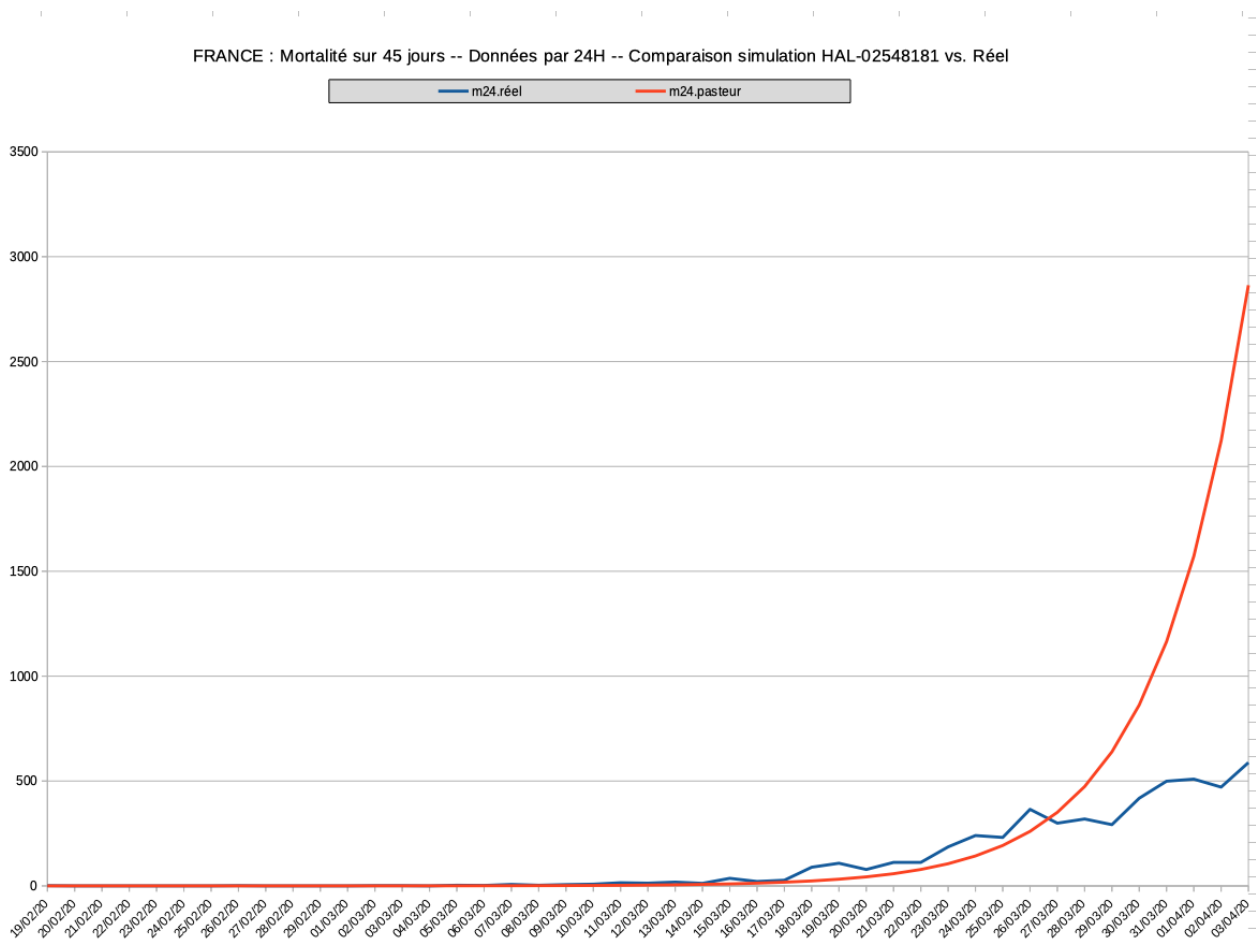
- données de mortalité issues des infections

Le niveau de cette mortalité se déduit du niveau des infections en appliquant le taux de mortalité de l'étude, à savoir 0,53 % aux valeurs de la simulation (Daily new infections).

2.6 La comparaison Simulation vs. Réel

Nous avons représenté sur un même graphique (Figure 1) les données de la simulation (échelle normale) et les données réelles.

Sur la période de 45 jours, l'adéquation entre la simulation de l'étude et le réel n'est que de **38,3%**.



L'adéquation est appréciée en fonction d'une erreur quadratique classique :

Adéq = $1 - \sqrt{\frac{\sum(Vr - Vm)^2}{\sum Vr}}$ avec Vr valeurs réelles et Vm valeurs du modèle.

3 Conclusion

La méthode de simulation utilisée consiste en une présentation des données en échelle logarithmique base 10 et droite de régression comme meilleure approximation.

Cette méthode est très simple mais elle n'est vraiment adaptée que quand les données réelles ont un comportement de type exponentiel pur, ce qui n'est visiblement pas le cas dans cette épidémie-ci pour ces données réelles pré-confinement.

Un tel lissage n'est pas condamnable en soi, mais s'il est utilisé comme instrument d'extrapolation, en particulier en « fin de courbe » au moment de la date d'impact du confinement, alors il devient inapproprié, pouvant conduire à des extrapolations douteuses.