

## Développement d'un modèle mécanistico-statistique pour calculer le taux de létalité réel du COVID-19

**Dès fin mars, une équipe INRAE a pu calculer le taux de létalité lié au COVID-19 pour la première fois au monde dans un pays autre que la Chine. Ils ont obtenu un taux de 0,5% sur la base des données hospitalières françaises, et de 0,8% en intégrant les données issues des EHPAD. Leurs résultats sont confirmés fin avril par les calculs de l'Institut Pasteur et une étude menée à New York. Leur étude vient d'être publiée le 8 mai 2020 dans la revue *MDPI Biology*.**

L'évolution d'une épidémie est toujours difficile à évaluer car elle dépend de processus latents que les chercheurs et médecins ne peuvent pas observer directement comme le processus de contagion du virus. Il est pourtant crucial de connaître au plus tôt le taux de létalité du virus et d'avoir une estimation du nombre de personnes infectées pour gérer au mieux une crise sanitaire. Au début d'une épidémie, le taux de létalité apparent (CFR, pour case fatality rate), correspondant au nombre de décès imputables à la maladie divisé par le nombre de cas connus, ne donne que peu d'information sur le risque associé à l'épidémie, car il est très dépendant du taux de dépistage. Par exemple, ce taux de létalité apparent est de 19,0% en France au 10 mai mais ne correspond pas à la réalité de l'épidémie. C'est donc le taux de létalité réel (IFR, pour infection fatality rate) correspondant au nombre de décès imputables au COVID-19 divisé par le nombre de cas réels, qu'il convient d'estimer. Au mois de mars, le calcul du taux de létalité réel du COVID-19 se basait uniquement sur les données issues de deux populations : les personnes rapatriées de Wuhan et les passagers du bateau de croisière Diamond Princess. Deux populations très spécifiques qui ne sont pas des échantillons représentatifs de la population globale.

En épidémiologie, il existe deux approches que les chercheurs peuvent utiliser pour calculer le taux de létalité, mais qui présentent chacune des inconvénients :

- L'approche statistique (data-driven) qui se base sur les données rassemblées et analysées par les chercheurs (nombre de personnes testées positives, nombre de décès...) mais qui ne prend pas en compte les mécanismes propres de l'épidémie (mode de contagion du virus, contacts entre les personnes...).
- L'approche mécanistico qui se base sur ces mécanismes, mais dont les sorties sont difficilement comparables aux données de terrain car celles-ci dépendent fortement de la façon de collecter les données (biais d'échantillonnage, erreurs d'observation, ...). Les données correspondent de fait à une observation très indirecte du processus, c'est pour cela qu'on parle de processus latent (ou caché).

Des chercheurs INRAE des départements scientifiques NUMM et SPE développent depuis plusieurs années des approches mécanistico-statistiques pour construire des modèles couplant les deux approches précédentes. S'ils travaillent essentiellement sur des épidémies végétales et des invasions biologiques, les principes développés restent les mêmes pour les épidémies animales et humaines. Ces méthodes sont particulièrement utiles lorsque l'observation (ici le nombre de cas testés positifs) ne correspond pas au processus lui-même, mais dépend indirectement de ce processus latent.

## Développement du modèle

Les approches mécanistico-statistiques sont particulièrement pertinentes pour l'épidémie de COVID-19 : si l'on dispose d'un grand nombre de données de terrain, celles-ci sont sujettes à de nombreux biais. Ici le processus latent est décrit grâce à un modèle épidémiologique standard qui prend en compte les caractéristiques propres au virus, et certains paramètres à estimer (dont le fameux  $R_0$ ). Un modèle probabiliste est utilisé pour faire le lien entre les données observées (nombre de cas connus, nombre de tests effectués, nombre de décès) et ce modèle épidémiologique. Ce modèle probabiliste d'observation comporte lui aussi un paramètre qui sera estimé : la probabilité relative de subir un test pour un individu « sain » vs un individu infecté par le SARS-CoV-2 (COVID-19).

Ces méthodes ont permis aux chercheurs de calculer, dès le mois de mars 2020, le taux de létalité lié à l'épidémie de COVID-19 pour la première fois dans un pays autre que la Chine, la France. Ils ont calculé un taux de létalité de 0,5% en se basant sur les données hospitalières et un taux de 0,8% en intégrant les données issues des EHPAD. Simultanément, ils ont calculé un taux de reproduction de base  $R_0$  de 3,2. Leurs résultats sont confirmés fin avril par les calculs de l'Institut Pasteur, qui trouve également un taux de 0,5% sur la base des données hospitalières françaises et par une étude sérologique menée à New-York où un taux de 0,6% est obtenu. Ces approches mécanistico-statistiques présentent un grand intérêt en épidémiologie car elles permettent de calculer très tôt dans l'avancée de l'épidémie des paramètres cruciaux comme le taux de létalité.

## Références

Roques, L.; Klein, E.K.; Papaix, J.; Sar, A.; Soubeyrand, S. *Using Early Data to Estimate the Actual Infection Fatality Ratio from COVID-19 in France*. MDPI Biology Mai 2020, 9, 97. <https://doi.org/10.3390/biology9050097>

Blog de l'équipe de Recherche BioSP

<https://informatique-mia.inra.fr/biosp/COVID-19>

### Contacts scientifiques :

Lionel Roques – [lionel.roques@inrae.fr](mailto:lionel.roques@inrae.fr)

Samuel Soubeyrand – [samuel.soubeyrand@inrae.fr](mailto:samuel.soubeyrand@inrae.fr)

Unité de recherche Biostatistique et processus spatiaux (BioSP)

Départements scientifiques NUMM et SPE

Centre INRAE Provence-Alpes-Côte d'Azur

### Contact presse :

Service de presse INRAE : 01 42 75 91 86 – [presse@inrae.fr](mailto:presse@inrae.fr)

---

INRAE, l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement, est un acteur majeur de la recherche et de l'innovation créé le 1er janvier 2020. Institut de recherche finalisé issu de la fusion entre l'Inra et Irstea, INRAE rassemble une communauté de 12 000 personnes, avec 268 unités de recherche, service et expérimentales implantées dans 18 centres sur toute la France. L'institut se positionne parmi les tout premiers organismes de recherche au monde en sciences agricoles et alimentaires, en sciences du végétal et de l'animal, et se classe 11ème mondial en écologie-environnement. Il est le premier organisme de recherche mondial spécialisé sur l'ensemble « agriculture-alimentation-environnement ». INRAE a pour ambition d'être un acteur clé des transitions nécessaires pour répondre aux grands enjeux mondiaux. Face à l'augmentation de la population, au changement climatique, à la raréfaction des ressources et au déclin de la biodiversité, l'institut construit des solutions pour des agricultures multi-performantes, une alimentation de qualité et une gestion durable des ressources et des écosystèmes.