



RÉPUBLIQUE FRANÇAISE



**OFFICE PARLEMENTAIRE D'ÉVALUATION
DES CHOIX SCIENTIFIQUES ET TECHNOLOGIQUES**

Note à l'attention des membres de l'Office

**Épidémie de COVID-19 – Point sur la modélisation épidémiologique pour estimer
l'ampleur et le devenir de l'épidémie de COVID-19**

Cette note a été présentée en réunion de l'Office parlementaire d'évaluation des choix scientifiques et technologiques le 30 avril 2020, par son premier vice-président, le député Cédric Villani, et a été mise à jour le 11 mai 2020 et validée pour publication.

L'épidémiologie est une branche interdisciplinaire de la science car elle utilise des concepts mathématiques, comme les statistiques, pour étudier des phénomènes biologiques, notamment les épidémies. C'est une science ancienne, qui s'est singulièrement développée avec les grandes épidémies telles que la grippe espagnole, puis, avec les possibilités de simulation numérique offertes par les progrès informatiques.

Modéliser une épidémie consiste à en créer une représentation mathématique. Différents cadres conceptuels existent pour de telles représentations, mais, une fois le cadre fixé, le modélisateur doit déterminer les valeurs des paramètres qui feront correspondre au mieux « l'épidémie modélisée » à l'épidémie réelle, pour la représentation choisie. La modélisation suppose donc des « allers-retours » entre calculs mathématiques et données issues du terrain (épidémiologie ou recherches médicales).

La modélisation est d'abord descriptive et explicative : elle permet de reproduire aussi fidèlement que possible les informations connues, passées ou présentes, donc de mieux connaître les facteurs gouvernant le développement de l'épidémie¹. Elle acquiert de ce fait une capacité prédictive, et devient donc aussi un instrument de la prise de décision. Cela peut servir plusieurs objectifs :

- Évaluation de l'incidence réelle de l'épidémie dans la population, ce qui permet de mieux estimer le niveau d'immunisation de la population, si l'on connaît le taux d'immunité sur des échantillons de population ;
- Appréciation des besoins en matière d'hospitalisation à court et moyen termes, pour mieux préparer l'afflux de patients à l'hôpital ;
- Prévision à court terme des besoins en ressources hospitalières pour adapter les mesures telles que le confinement, la fermeture des commerces, etc. ;
- Estimation, a priori, et évaluation, a posteriori, de l'intérêt des différentes mesures dans la limitation de la propagation de l'épidémie.

¹ http://covid-ete.ouvaton.org/Rapport7_resume.html

Secrétariats de l'OPECST :

Assemblée nationale : 01 40 63 26 81 – Sénat : 01 42 34 31 07

La présente note se propose d'examiner les fondements de la modélisation épidémiologique et d'examiner les prévisions qu'elle permet d'établir concernant l'épidémie de COVID-19.

Les modèles compartimentaux

Le modèle SIR, très simple, permet déjà d'éclairer correctement les phénomènes épidémiques

Les modèles épidémiologiques compartimentaux dérivent des travaux de Kermack et McKendrick sur l'épidémie de peste en Inde², eux-mêmes dérivés du modèle dit « proie-prédateur », mis en équation par Lotka et Volterra au XX^e siècle³. Ils sont utilisés pour décrire l'évolution temporelle d'une maladie au sein d'une population. En première approche, cette méthode répartit les individus dans trois catégories, représentant leur état face à la maladie et appelées *compartiments*. Dans les modèles dits SIR on distingue les compartiments :

- S, dans lequel se trouvent les individus sains et susceptibles d'être infectés (*susceptibles*) ;
- I, pour les individus infectés et infectieux (*infected*) ;
- R, pour les individus retirés du modèle (*removed*), non susceptibles d'être infectés, car guéris et immunisés ou décédés⁴.

Le modèle SIR décrit l'évolution dans le temps du nombre de personnes dans chaque compartiment et part du schéma simplifié suivant : un individu est d'abord sain (S) ; il est infecté (I) avec une probabilité β , représentant le taux de transmission de la maladie d'une personne infectée à une personne saine ; enfin, il guérit (R), avec la probabilité γ , représentant le taux de guérison, c'est-à-dire l'inverse de la durée moyenne des symptômes.



On comprend intuitivement que :

- pendant toute l'épidémie, l'effectif des personnes saines et susceptibles d'être infectées (S – en bleu) diminue continûment et l'effectif de personnes guéries (R – en rouge) augmente continûment ; en effet, par construction, il n'existe pas de flux d'entrée vers S et pas de flux de sortie depuis R ;

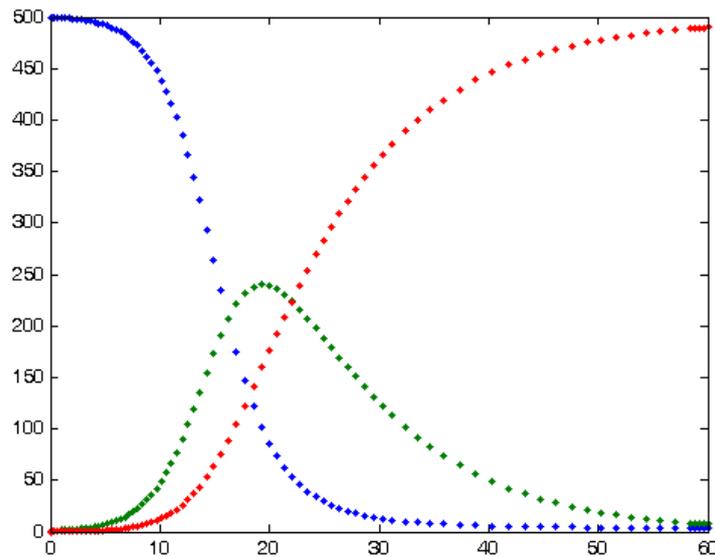
2 Kermack, McKendrick, et Walker. « A contribution to the mathematical theory of epidemics ». *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115, n° 772 (1 août 1927): 700-721. <https://doi.org/10.1098/rspa.1927.0118>.

3 Celui-ci décrit la régulation naturelle qui s'opère entre les populations animales, relativement aux chaînes alimentaires. L'effectif de chaque population est gouverné par des facteurs qui lui sont propres (notamment le taux de natalité) et par un couplage avec la population d'une autre espèce (consommation de l'espèce « proie » par l'espèce « prédateur »). Il s'agit d'un modèle déterministe qui peut s'appliquer aux épidémies, en première approche.

4 Le modèle SIR le plus simple considère que le virus ne tue pas et que la population totale est donc constante. Même pour l'épidémie de COVID-19, cette approximation est valable à l'échelle de la population française, le nombre de décès étant inférieur de plusieurs ordres de grandeur à la taille de la population.

– lorsque l'épidémie débute, le nombre des personnes infectées (I – en vert) augmente rapidement : en effet, le virus trouve très facilement des hôtes à infecter dans une population très majoritairement saine ;

– mais, puisque la population susceptible d'être infectée décroît et que la population qui ne peut plus l'être s'accroît, la dynamique de l'infection s'affaiblit progressivement ; au bout d'un temps suffisamment long, quand il y a un très grand nombre de personnes guéries (R) et très peu de personnes susceptibles (S), le nombre de personnes infectées (I) ne peut être que faible lui aussi, et décroître.



Évolution du nombre de personnes dans chaque compartiment en fonction du temps, dans l'hypothèse où tous les individus sauf un sont sains au début de la simulation (S : bleu, I : vert, R : rouge)

L'évolution temporelle du nombre d'individus sains est régie par les équations différentielles suivantes :

$$\frac{dS}{dt} = -\beta S(t)I(t)$$

$$\frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t)$$

$$\frac{dR}{dt} = \gamma I(t)$$

S(t), I(t) et R(t) représentant respectivement le nombres d'individus sains, infectés et retirés à l'instant t. L'on peut noter que la somme des sous-population est constante, elle est égale à la taille de la population totale :

$$N = S(t) + I(t) + R(t)$$

L'évolution du nombre cumulé de personnes infectées connaît **une dynamique exponentielle dans les débuts de l'épidémie**, jusqu'à ce qu'environ la moitié de la population totale ait été infectée. Cette croissance est très rapide jusqu'à ce que la raréfaction des individus sains, l'accroissement de la part des individus immunisés, ainsi que, potentiellement, les mesures prises contre l'épidémie, viennent ralentir considérablement son évolution.

Par ailleurs, on montre qu'il existe un moment où le nombre de personnes infectées I atteint un maximum (voir courbe verte dans le graphique ci-dessus), qui peut être calculé par simulation numérique. Cela signifie que **la pression sur le système de santé passe elle aussi par un maximum**. À ce degré de simplicité, la modélisation ne permet pas d'en dire plus sur les conséquences de l'épidémie, car la gravité des différentes formes de la maladie n'est pas prise en compte. Le modèle SIR le plus simple montre cependant qu'il est pertinent de mettre en regard deux maximums : le nombre maximum de personnes infectées et la capacité maximale du système de santé à prendre en charge les patients.

Le nombre de reproduction, un paramètre déterminant dans l'évolution de l'épidémie

• Le paramètre β est spécifique de la maladie et représente le taux d'infection, c'est-à-dire la probabilité qu'un contact entre un individu infecté et un individu sain provoque l'infection de ce dernier. Il peut être écrit comme un rapport ($\beta = \frac{R_0}{T}$), où T est la durée moyenne de contagiosité et R_0 , appelé « nombre de reproduction de base » est le nombre moyen d'individus sains que peut contaminer une personne infectée pendant le temps T où elle est contagieuse. Si la notion est relativement simple à comprendre, le calcul ou l'estimation de ce nombre s'avère plus complexe.

On comprend que :

- Si $R_0 < 1$: un individu (I) contamine en moyenne moins d'une personne (S), l'épidémie ne peut pas s'installer ou, si elle est déjà installée, elle s'éteint rapidement ;
- Si $R_0 > 1$: un individu (I) contamine en moyenne plus d'une personne (S), l'infection est envahissante et dégénère en épidémie.

Ce paramètre permet de mettre en évidence un seuil d'« immunité grégaire » ou « immunité de groupe », au-dessus duquel la population est suffisamment immunisée pour que tout foyer de l'épidémie tende à s'éteindre :

$$\text{Seuil d'immunité grégaire} = 1 - \frac{1}{R_0}$$

Il convient de noter que ce n'est pas parce que la fraction des personnes « retirées » devient supérieure au seuil d'immunité de groupe que l'épidémie cesse de s'étendre : elle a certes perdu sa dynamique exponentielle mais, tout en s'éteignant peu à peu, elle continue à infecter de nouvelles personnes.

Le calcul de la fraction de la population qui, en fin d'épidémie, n'a jamais rencontré le virus, et qui reste donc dans le compartiment S des personnes saines, fait également intervenir R_0 . Plus ce paramètre est élevé, plus cette fraction de la population est faible. Les simulations numériques montrent qu'elle est d'environ 5 % lorsque R_0 est égal à 3.

Dans le cadre de l'épidémie de COVID-19, dans l'hypothèse la plus probable où le nombre de reproduction de base se situerait entre 2 et 3, le seuil d'immunité de groupe serait de 50 à 67 % de la population. Dans le cas de la rougeole⁵, le nombre de reproduction de base est d'environ

5 JR Zahar, MD, GHU Necker – Enfants Malades, Université Paris Descartes, « Rougeoles et Coqueluches nosocomiales », 2012 (<https://www.infectiologie.com/UserFiles/File/medias/JNI/JNI12/2012-JNI-Rougeole-Coqueluche-noso-zahar.pdf>).

18, ce qui conduit à un seuil d'immunité grégaire de 95 % ; pour la coqueluche⁶, le nombre de reproduction de base est d'environ 16, et le seuil d'immunité grégaire est ainsi de 94 %.

Le concept de « seuil d'immunité grégaire » peut viser tant l'immunité acquise naturellement au cours de l'épidémie que l'immunité acquise volontairement par le biais de la **vaccination**. Dans ce deuxième cas, les pourcentages mentionnés précédemment doivent être lus comme la proportion minimale de la population qu'il convient de vacciner pour que le nombre de reproduction de base soit inférieur à 1 et qu'en conséquence, l'infection ne dégénère pas en épidémie.

Bien qu'extrêmement simple, le modèle SIR donne donc un fondement rigoureux aux politiques de vaccination. Il a aidé à mettre en place la politique de vaccination contre la variole au début du XX^e siècle⁷.

• Si le nombre de reproduction de base R_0 caractérise l'évolution de l'épidémie dans une population lorsqu'aucune mesure n'est prise afin de la limiter, le nombre de reproduction effectif (R_{eff}) décrit son évolution sous les contraintes en vigueur. R_{eff} peut être défini par le produit du nombre de reproduction de base et de la fraction restante de la population susceptible d'être infectée. On peut approximativement définir R_{eff} ainsi :

$$R_{eff} = m \times (1 - p) \times \frac{\beta}{\gamma}$$

Avec m la fraction de la population susceptible d'être infectée, p la probabilité qu'un individu infecté soit isolé, β le taux de transmission et γ le taux de guérison, R_0 pouvant être approximé par le rapport du taux de transmission et du taux de guérison (β/γ).

Pour ralentir l'épidémie, il est essentiel de diminuer R_{eff} autant que possible et d'essayer le faire passer au-dessous de 1. Sa définition met en exergue plusieurs stratégies qui peuvent être utilisées de concert :

- Diminuer la fraction de la population susceptible d'être infectée en réduisant le nombre de contacts de chaque individu ;
- Diminuer la probabilité de transmission, lors d'un contact ;
- Augmenter l'isolement des cas d'infection.

La modélisation épidémiologique peut être utilisée a posteriori, pour évaluer l'impact de mesures, telles que le port du masque, sur la propagation de l'épidémie. Dans les faits, beaucoup des paramètres utilisés pour alimenter les modèles proviennent des observations réalisées en Chine. Il n'est en général pas possible d'explicitier l'impact d'une mesure déterminée : le plus souvent, la modélisation peut seulement apprécier l'impact de l'ensemble des mesures prises pour ralentir la propagation de l'épidémie.

Alors que les estimations du nombre de reproduction de base de l'épidémie variaient⁸ de 2 à 4, des scientifiques estimaient au début du mois de mars que le traçage des cas et leur isolement

6 Ibid.

7 Corentin Bayette et Marc Monticelli, « Crise sanitaire du Coronavirus Covid-19. Modélisation d'une épidémie. Comment les maths aident à la prise de décision », CNRS – Images des mathématiques (12 avril 2020) (<https://images.math.cnrs.fr/Modelisation-d-une-epidemie-partie-1.html>).

8 Majumder et Mandl. « Early in the Epidemic: Impact of Preprints on Global Discourse about COVID-19 Transmissibility ». *The Lancet Global Health* (24 mars 2020). [https://doi.org/10.1016/S2214-109X\(20\)30113-3](https://doi.org/10.1016/S2214-109X(20)30113-3).

sous la forme de quatorzaine abaisserait le R_{eff} à 1,5 en Chine – à l'exclusion de la province de Hubei – ce qui est insuffisant pour que l'épidémie s'éteigne d'elle-même⁹. En Corée du Sud, le R_{eff} estimé à la fin du mois de février était de 1,5 également¹⁰. Fin mars, une autre étude sud-coréenne l'estimait toujours compris entre 1 et 2 dans les régions qui n'étaient pas sièges des principaux foyers de l'épidémie dans le pays¹¹.

Comme indiqué précédemment, un nombre de reproduction supérieur à 1 est synonyme d'une épidémie prenant de l'ampleur. Un pilotage fin de la réponse sanitaire nécessite donc d'évaluer régulièrement et de la façon la plus précise possible l'évolution de ce paramètre. Cette nécessité a été largement intégrée et reprise par les gouvernements allemands et français, signe que les scientifiques sont très écoutés dans cette crise. L'exemple sud-coréen est signe que même si un pays ne présente que peu de cas, l'épidémie peut tout de même suivre une dynamique de croissance.

L'intervalle de génération, ou intervalle sériel, un autre paramètre déterminant

L'intervalle de génération se définit comme le temps moyen entre deux contaminations dans une chaîne de transmission, c'est-à-dire le temps qu'il faut pour qu'une personne nouvellement infectée en infecte une autre. Plus cet intervalle est court, plus la dynamique de l'épidémie est rapide.

Cet intervalle est déterminé par l'étude de chaînes de transmission bien documentées. Il a été estimé à 4 jours pour le COVID-19¹², alors qu'il était de 8 jours pour le SRAS.

Les inconvénients du modèle SIR simple

Le modèle considère que les interactions sociales sont aléatoires – chaque personne est supposée avoir la même probabilité de contacter n'importe quelle autre – et que les populations sont homogènes – tous les individus ont un comportement semblable. De ce fait, le modèle SIR et, plus généralement, les modèles compartimentaux peuvent mésestimer la complexité des interactions entre individus. Les connexions sociales, au-delà du domicile, de l'école et du lieu de travail, sont souvent très structurées. Certains individus ont des dizaines de contacts alors que d'autres n'en ont que très peu, et les contacts de chaque individu forment eux-mêmes des structures complexes. La considération de ces différences, notamment par tranche d'âge, est nécessaire, car surestimer l'importance des interactions dans la propagation de la maladie, pourrait mener à surestimer l'efficacité des mesures visant à réduire les interactions entre individus.

Des simulations prenant en compte le fait que les populations se regroupent en sous populations distinctes, plus ou moins homogènes, et plus ou moins interconnectées, existent. Au sein de

9 Li, *et al.* « Transmission Characteristics of the COVID-19 Outbreak in China: A Study Driven by Data ». *MedRxiv*, 1 mars 2020. <https://doi.org/10.1101/2020.02.26.20028431>.

10 Shim, *et al.* « Transmission Potential and Severity of COVID-19 in South Korea ». *International Journal of Infectious Diseases* 93 (1 avril 2020): 339-44. <https://doi.org/10.1016/j.ijid.2020.03.031>.

11 Hwang, *et al.* « Basic and Effective Reproduction Numbers of COVID-19 Cases in South Korea Excluding Sincheonji Cases ». *MedRxiv*, 31 mars 2020. <https://doi.org/10.1101/2020.03.19.20039347>.

12 Nishiura *et al.* « Serial Interval of Novel Coronavirus (COVID-19) Infections ». *International Journal of Infectious Diseases* 93 (1 avril 2020): 284-86. <https://doi.org/10.1016/j.ijid.2020.02.060>.

chaque sous-population, une approche compartimentale est appliquée, et une approche en réseau pour relier les sous-populations entre elles¹³.

D'autre part, la probabilité de transmission dépend du comportement de l'individu, comme l'utilisation plus ou moins fréquente des transports en commun, ou l'exercice d'une activité régulière dans un espace confiné. Ces pratiques dépendent fortement du milieu de vie, urbain ou rural, et du mode de vie.

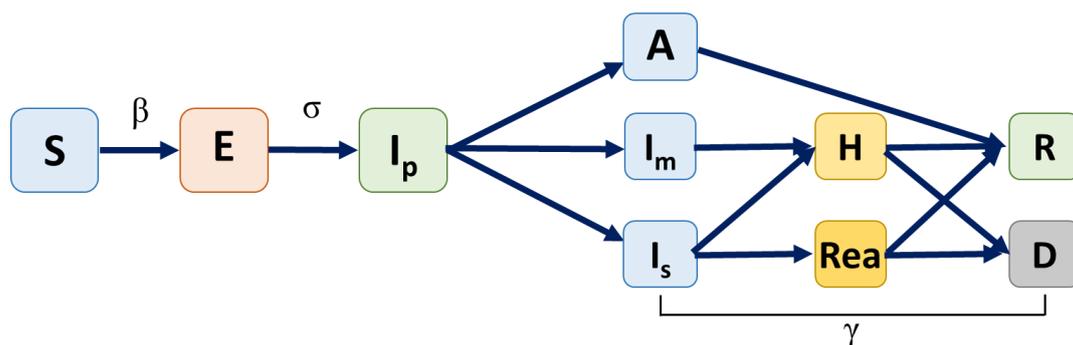
Complexifier pour affiner

Une description plus précise de la population par ajout de compartiments supplémentaires

Le modèle SIR peut être complexifié par l'ajout de compartiments supplémentaires, qui permettent de décrire de façon plus précise la population au regard de ses caractéristiques vis-à-vis des processus d'infection et de guérison.

Dans le cadre de l'épidémie de COVID-19, et à la différence de l'épidémie de SRAS, un compartiment E est nécessaire pour caractériser les individus exposés au virus mais pas encore symptomatiques, ainsi qu'un compartiment I_p pour caractériser l'intervalle de temps au cours duquel les personnes sont contagieuses, avant l'arrivée des symptômes¹⁴.

On peut également ajouter des compartiments dans des branches parallèles permettant de différencier les cas sévères, qui occupent un lit d'hôpital, des cas moins graves. Cette distinction est particulièrement nécessaire lorsque la modélisation a pour objectif de prédire les besoins en hospitalisation. Il faut noter qu'une telle démarche, qui concerne « l'aval » de l'infection proprement dite, n'a pas de conséquences sur la modélisation de la dynamique épidémique à proprement parler, qui se situe en amont. Elle permet d'obtenir des informations utiles aux autorités sur la charge susceptible de peser sur les différents secteurs composant le système de santé.



Exemple de modèle

S : susceptible, E : exposé, I_p : infecté présymptomatique, A : asymptomatique, I_m : infecté symptômes moyens, I_s : infecté, symptômes sévères, H : admis à l'hôpital, Rea : admis en réanimation, R : guéri et

13 M. Chinazzi et al., « The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak », Science, mars 2020, doi: 10.1126/science.aba9757.

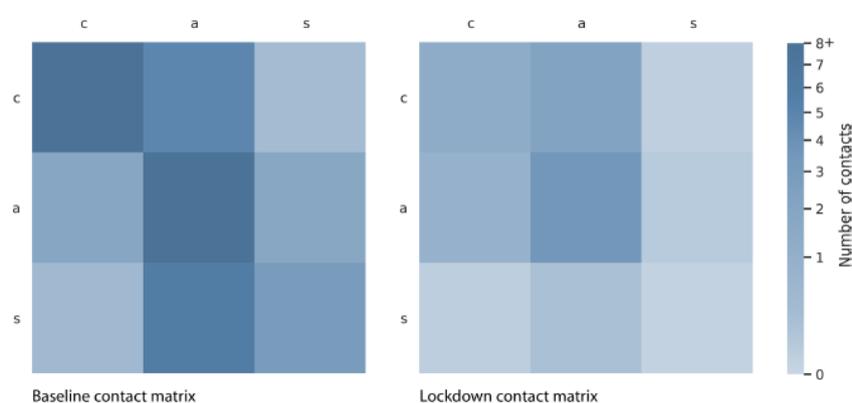
14 Lescure, et al. « Clinical and Virological Data of the First Cases of COVID-19 in Europe: A Case Series ». *The Lancet Infectious Diseases*, 27 mars 2020. [https://doi.org/10.1016/S1473-3099\(20\)30200-0](https://doi.org/10.1016/S1473-3099(20)30200-0).

D : décédé ; β : taux de transmission de la maladie, σ : taux d'incubation (1/ durée d'incubation), γ : taux de guérison (1/ durée de la maladie).

Une description plus précise des interactions sociales grâce à des « matrices de contact »

Des études se sont intéressées à la nature et la fréquence des contacts dans la population, de façon à élaborer des matrices de contact. Celles-ci permettent, par exemple, de visualiser le fait que les interactions intra et inter classes d'âge sont différentes. La prise en compte de l'hétérogénéité des interactions dans la population est alors possible par une stratification du modèle en fonction de l'âge¹⁵ qui rend le modèle plus réaliste.

De telles matrices peuvent être réalisées dans différents contextes : le foyer, le travail, l'école. Ainsi, la matrice « école » peut être retirée du modèle pour examiner l'effet de la fermeture des établissements scolaires sur la dynamique de l'épidémie. Ce niveau de réalisme est donc particulièrement utile pour examiner les stratégies de sortie de confinement.



Exemple de matrice de contact entre les enfants (c), les adultes (a) et les seniors (s), en temps normal (gauche) et dans un contexte de confinement (droite)¹⁶.

Une stratification de la population par âge a également un intérêt dans la mesure où la probabilité de développer des formes graves varie avec l'âge¹⁷.

D'autres types de complexification

En France, il y a un intérêt à considérer une certaine hétérogénéité spatiale, dans la mesure où le confinement a bloqué l'épidémie à des stades assez différents d'une région à l'autre, les régions de l'Ouest de la France métropolitaine étant moins affectées que les régions de l'Est. Des modèles affinés au niveau régional sont nécessaires, notamment pour prendre en compte les différences en termes de capacités hospitalières et de démographie. On observe également de grandes différences de prévalence de la maladie au niveau subrégional, comme entre les départements d'Île-de-France. La compréhension de ces différences, si elles reposent sur d'autres paramètres que le simple nombre de contacts, doit permettre d'affiner les modèles.

15 Domenico, *et al.* « Expected Impact of Lockdown in Île-de-France and Possible Exit Strategies », pré-publication mise en ligne le 12 avril 2020.

16 Domenico, *et al.* cf note 10.

17 « COVID-19, Point épidémiologique hebdomadaire du 16 avril 2020 » de Santé publique France ; <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-16-avril-2020>

Alors que la résolution des équations différentielles du modèle SIR implique un traitement de la simulation en temps continu, il est possible de discrétiser le modèle, c'est-à-dire de faire avancer le modèle étape par étape, de façon à pouvoir incorporer un effet mémoire. Celui-ci est plus réaliste dans la mesure où il permet de considérer qu'un individu entré dans le compartiment I a une probabilité d'autant plus grande d'en sortir qu'il y est entré depuis longtemps, tandis qu'un individu qui vient d'y rentrer a une probabilité plus faible d'en sortir dans l'immédiat¹⁸. Il est néanmoins possible de traduire cet effet en utilisant une distribution exponentielle des variables aléatoires, au sein de chaque compartiment, dans une simulation en temps continu. Une telle démarche permet de prendre en compte des phénomènes qui sont masqués par le fait que, dans le modèle SIR classique, la population de chaque compartiment est considérée comme homogène. Par exemple, cette démarche permettra de modéliser les conséquences d'une hypothèse selon laquelle un individu infecté (compartiment I) serait plus contagieux au début qu'à la fin de la période d'infection – dans le modèle SIR classique, il est « également contagieux » pendant ladite période.

Avantages et inconvénients

Les modèles les plus réalistes ne sont pas nécessairement les plus fiables sur le long terme. Cependant, ils sont utiles lorsqu'il s'agit d'estimer si un paramètre a une grande influence sur le devenir à court et moyen terme de l'épidémie, pour mieux envisager une sortie de confinement, par exemple.

La fiabilité des prédictions des modèles dépend de l'exactitude des paramètres qui les alimentent, avec différentes sensibilités, alors que leur estimation évolue au fur et à mesure de l'avancée des connaissances sur l'épidémie. L'inconvénient majeur des modèles les plus complexes est leur dépendance à un grand nombre de paramètres, qui ne sont pas tous estimés avec une grande précision¹⁹. Ces paramètres viennent d'études empiriques, fiables, telles que celles réalisées en milieu fermé, comme sur le navire de croisière « Princess Diamond »²⁰, des données d'hospitalisation et de décès, ou bien d'études réalisées sur des échantillons de population. Pour estimer la prévalence de l'infection, une étude réalisée dans la ville de New-York sur les femmes enceintes admises pour un accouchement a ainsi montré que 15 % d'entre elles étaient porteuses du virus²¹. Ces observations empiriques sont néanmoins souvent biaisées : la moyenne d'âge était relativement élevée sur le navire de croisière, et les conditions de vie à bord étaient probablement différentes de la vie réelle, par exemple. Il est nécessaire de croiser les informations obtenues par différents moyens d'observation afin de limiter ces biais.

18 http://covid-ete.ouvaton.org/Rapport7_resume.html

19 Par exemple, Djidjou-Demasse et collaborateurs ont différencié les cas d'individus infectés avec des symptômes moyens de ceux présentant des symptômes sévères. Ces derniers étant plus susceptibles d'être hospitalisés, ou, à tout le moins, identifiés comme malades, ils sont probablement moins contagieux que des individus infectés dont l'entourage en se méfierait pas. Pour inclure ce paramètre dans le modèle, un paramètre « taux de réduction de la contagiosité des infections sévères » a été estimé à 0,2 ; c'est-à-dire que l'individu présentant des symptômes sévères est contagieux à hauteur de 20 % de la contagiosité des individus présentant des symptômes modérés ; Djidjou-Demasse, *et al.* « Optimal COVID-19 Epidemic Control until Vaccine Deployment ». *MedRxiv*, 6 avril. <https://doi.org/10.1101/2020.04.02.20049189>.

20 Une estimation de la part des cas asymptomatiques a été réalisée à partir de l'étude des passagers du navire Princess Diamond : entre 15 et 20 % des personnes infectées seraient asymptomatiques. Cette estimation doit être étendue à la population générale avec précaution, car le navire ne représente pas un échantillon représentatif de la population ; Mizumoto, *et al.* « Estimating the Asymptomatic Proportion of Coronavirus Disease 2019 (COVID-19) Cases on Board the Diamond Princess Cruise Ship, Yokohama, Japan, 2020 ». *Eurosurveillance* 25, n° 10 (12 mars 2020): 2000180. <https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180>.

21 Sutton, *et al.* « Universal Screening for SARS-CoV-2 in Women Admitted for Delivery ». *New England Journal of Medicine* (13 avril 2020). <https://doi.org/10.1056/NEJMc2009316>.

Certaines nouvelles connaissances nécessitent parfois même la création d'un nouveau compartiment, et pas seulement l'ajustement d'une probabilité. Cela a été le cas lorsqu'on a montré que les personnes asymptomatiques infectées peuvent être contagieuses. Il reste beaucoup d'inconnues quant à l'immunité – acquise et existante – à l'égard de ce virus. D'une part, il n'est pas encore certain que tous les individus (R) soient immunisés, donc qu'ils soient définitivement écartés de la population susceptible d'être infectée (S), d'autre part, les modèles partent tous du principe qu'il n'y a aucun individu immunisé à l'origine, alors qu'une immunité croisée est possible avec les coronavirus responsables de rhumes²² circulant en Europe.

Finalement, alors que l'objectif premier des modèles est de prévoir l'évolution de la maladie dans un contexte précis, avec des mesures visant à limiter la propagation de la maladie, la robustesse de leurs conclusions est affectée par l'incertitude relative à l'adoption effective des gestes barrières ou des mesures de distanciation physique par la population²³.

Résultats des simulations pour la France

Quelle est l'ampleur réelle de l'épidémie ?

Sur la base de telles simulations, des estimations de l'ampleur réelle de l'épidémie ont été proposées. Une étude britannique dirigée par Neil Ferguson et ses collaborateurs estimait que 3 % (entre 1 % et 7 %) de la population française avait été infectée par le virus à la date du 28 mars²⁴. L'équipe de modélisateurs réunie autour de Simon Cauchemez, à l'Institut Pasteur, indiquait pour sa part qu'entre 2 et 7 millions de Français devraient avoir contracté la maladie d'ici au 11 mai, soit un peu moins de 6 % de la population, loin du seuil d'immunité de groupe (de 50 à 66 %)²⁵.

Les travaux de Vittoria Colizza et de son équipe de l'Inserm se sont concentrés sur la région Île-de-France ; ils ont conclu qu'entre 1 % et 6 % de la population francilienne avait contracté le virus à la date du 5 avril²⁶.

L'équipe de l'Institut Pasteur s'est attachée à estimer le taux d'hospitalisation et de létalité pour la France : 2,6 % des personnes infectées seraient hospitalisées et le taux de létalité serait de 0,53 %²⁷.

22 Adam, David. « Special Report: The Simulations Driving the World's Response to COVID-19 ». *Nature* 580, n° 7803 (2 avril 2020): 316-18. <https://doi.org/10.1038/d41586-020-01003-6> et Stéphane Korsia-Meffre. « Persistance et efficacité des anticorps neutralisants contre le SARS-CoV-2 : état des connaissances et leçons des autres coronavirus humains ». VIDAL. Consulté le 22 avril 2020. https://www.vidal.fr/actualites/24770/persistance_et_efficacite_des_anticorps_neutralisants_contre_le_sars_cov_2_etat_des_connaissances_et_lecons_des_autres_coronavirus_humains/

23 Adam, David. « Special Report: The Simulations Driving the World's Response to COVID-19 ». *Nature* 580, n° 7803 (2 avril 2020): 316-18. <https://doi.org/10.1038/d41586-020-01003-6>.

24 Flaxman, *et al.* « Estimating the Number of Infections and the Impact of Non- Pharmaceutical Interventions on COVID-19 in 11 European Countries », 2020, 35.

25 Salje, *et al.* « Estimating the Burden of SARS-CoV-2 in France », pré-publication mise en ligne le 20 avril 2020 ; <https://hal-pasteur.archives-ouvertes.fr/pasteur-02548181/document>.

26 Domenico, *et al.* « Expected Impact of Lockdown in Île-de-France and Possible Exit Strategies », pré-publication mise en ligne le 12 avril 2020.

27 Salje, *et al.* « Estimating the Burden of SARS-CoV-2 in France », pré-publication mise en ligne le 20 avril 2020 ; <https://hal-pasteur.archives-ouvertes.fr/pasteur-02548181/document>.

Quel impact du confinement ?

L'équipe de modélisateurs réunie autour de Pascal Crépey, à l'EHESP, a montré que plus de 14 millions de Français auraient pu contracter la maladie en date du 19 avril si le confinement n'avait pas été mis en place, soit 20 % de la population, contre seulement 6 % estimés à cette même date, grâce au confinement. Ils ont également mis en évidence que près de 105 000 lits de réanimation auraient été nécessaires pour prendre en charge les malades présentant des formes graves, alors que le pays disposait de 5 000 lits de réanimation avant l'épidémie, du double maintenant. Ainsi, plus de 60 000 décès à l'hôpital auraient été évités²⁸.

L'équipe réunie autour de Simon Cauchemez a estimé que le confinement a réduit le nombre de reproduction effectif de la maladie de 3,3 à 0,5, quand celle de Vittoria Colizza l'a estimé à 0,7, sur la base d'une réduction globale des contacts de 80 %.

Dans certains pays européens n'ayant pas opté pour un confinement, tels que la Suède, on peut tout de même constater un abaissement du nombre de reproductif effectif sous la valeur seuil de 1²⁹.

Quel contrôle de l'épidémie après le déconfinement ?

Si le confinement visait à faire tomber le R_{eff} sensiblement au-dessous de 1 pour réduire au maximum la propagation du virus, ramener l'épidémie sous contrôle – gagner du temps afin de faire retomber la pression exercée sur le système de soins, notamment l'hôpital – le déconfinement doit viser, surtout en l'absence de vaccin, à maintenir le R_{eff} légèrement au-dessous de 1, par le biais de mesures variées – masques, détection et suivi des cas contacts, isolement, etc. – dont la portée peut devoir être ajustée en fonction des situations épidémiques rencontrées sur le territoire.

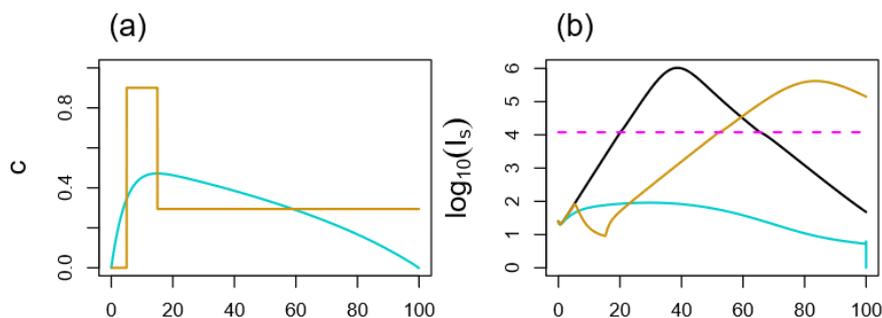
Le confinement n'ayant pas complètement stoppé l'épidémie, de nouvelles contaminations ont encore lieu. L'équipe de l'Institut Pasteur a estimé que début mai, il y aurait toujours 1 300 nouvelles contaminations par jour³⁰. Ainsi, les modèles suggèrent que si les mesures adoptées en France étaient soudainement abandonnées, l'épidémie reprendrait de plus belle.

Mais ces mesures, visant à réduire la probabilité de transmission par plusieurs biais, sont réellement efficaces, et leur maintien partiel pourrait limiter le nombre de nouveaux cas en dessous d'un seuil acceptable, correspondant aux capacités hospitalières. L'équipe réunie autour de Samuel Alizon a récemment démontré qu'un niveau de contrôle homogène et de long terme, calculé d'après la théorie du contrôle, serait plus efficace pour limiter l'incidence de l'épidémie qu'une réponse sanitaire en « tout ou rien », telle un confinement suivi d'une absence totale de contre-mesures.

28 Roux, Massonnaud et Crépey. « COVID-19: One-Month Impact of the French Lockdown on the Epidemic Burden », Pré-publication mise en ligne le 22 avril 2020 ; <https://www.ea-reperes.com/wp-content/uploads/2020/04/ImpactConfinement-EHESP-20200322v1.pdf>.

29 Le ministère de la santé suédois l'a estimé inférieur à 1 depuis le 20 avril ; <https://www.folkhalsomyndigheten.se/contentassets/4b4dd8c7e15d48d2be744248794d1438/sweden-estimate-of-the-effective-reproduction-number.pdf>

30 En moyenne, les nouvelles infections ont été estimées à 1 300 par jour, avec un intervalle de confiance de 840 à 2 300 nouvelles infections. Voir supra Salje, *et al.*



Modélisation de l'impact des stratégies de contrôle (à gauche : en jaune, un confinement puis un relâchement du contrôle ; en bleu, un contrôle optimal) sur l'incidence de la maladie (à droite, représentation logarithmique du nombre de cas d'infection dans le temps : en noir, absence de contrôle).

Le contrôle optimal, même s'il est moins intense que la stratégie de confinement, permet de limiter le nombre de cas d'infections sur le long terme, tandis que le relâchement soudain des mesures retarde l'apparition du pic mais ne l'empêche pas³¹.

Pour permettre une reprise de l'activité économique tout en maintenant l'épidémie à un niveau ne menaçant pas le système hospitalier, des mesures adéquates devront être prises. Parmi celles-ci, un traçage des chaînes de transmission, le plus exhaustif possible, devra être réalisé, couplé avec une stratégie de dépistage ciblée pour éradiquer tout foyer épidémique : « *Ce n'est pas sur la base d'un dépistage massif que nous mettrons fin au confinement mais c'est sur la base d'un dépistage massif que nous pourrions être amené à accroître notre surveillance une fois le confinement levé* »³².

Quelle évolution à long terme ?

Les épidémiologistes, se basant sur les épidémies de SARS et de MERS, ainsi que sur les coronavirus circulant en Europe, estiment que le SARS-CoV-2 devrait persister quelques années encore. Une étude américaine prévoit qu'il sera nécessaire de maintenir des mesures de distanciation jusqu'à l'horizon 2022, si aucun vaccin n'était disponible en grande quantité d'ici-là³³.

Les données manquent encore pour mieux apprécier son évolution sur le long terme, notamment quant à l'immunité que les personnes infectées développent, ou encore à la question de la saisonnalité.

31 Djidjou-Demasse, *et al.* « Optimal COVID-19 Epidemic Control until Vaccine Deployment ». *MedRxiv*, 6 avril 2020. <https://doi.org/10.1101/2020.04.02.20049189>.

32 Propos du Dr. Pierre-Louis Druais, membre du Conseil scientifique COVID-19 auprès du Gouvernement. Le Généraliste. « Coronavirus : dépister plus, mais pour quoi faire ? » Consulté le 2 avril 2020. <https://www.legeneraliste.fr/actualites/article/2020/03/29/coronavirus-depister-plus-mais-pour-quoi-faire-322053>.

33 Kissler, *et al.* « Projecting the Transmission Dynamics of SARS-CoV-2 through the Postpandemic Period ». *Science*, 14 avril 2020. <https://doi.org/10.1126/science.abb5793>.

Conclusions

La modélisation épidémiologique est un outil prédictif qui connaît des limites : structurelles, puisqu'un modèle est une représentation simplifiée de la réalité ; circonstancielle, du fait de la connaissance encore partielle de la maladie. Certaines prédictions, établies au Royaume-Uni, comme celle de l'incidence de la grippe aviaire en 2005, se sont révélées finalement très éloignées de la réalité, alors que d'autres, établies dans le cadre de l'épidémie de maladie de la vache folle dans les années 90 et fièvre aphteuse en 2001 se sont avérées justes³⁴. Néanmoins, les modèles constituent une réelle aide à la décision, tout particulièrement dans la recherche d'une stratégie de déconfinement, dans la mesure où ils permettent de tester l'influence de certaines décisions possibles, comme le déconfinement de certaines classes d'âge uniquement, ou encore la réouverture des écoles.

Il est important d'avoir des données robustes. Les données hospitalières sont précieuses, mais elles ne suffisent pas et des enquêtes poussées, sur de petits échantillons de population, sont nécessaires. En France, les capacités de dépistage ont beaucoup évolué depuis le début de l'épidémie, elles ne constituent donc pas une base fiable pour déterminer un taux d'incidence réaliste de l'épidémie à un moment donné. Les études séro-épidémiologiques, dont celle réalisée par l'Institut Pasteur pour mesurer le taux d'attaque autour du cluster du lycée de Crépy-en-Valois³⁵, sont utiles pour ajuster les modèles.

Le devenir de l'épidémie à long terme reste incertain, mais tous les modèles mettent en garde contre un relâchement brutal des mesures visant à diminuer la propagation de la maladie.

Finalement, si les conséquences du confinement sur l'incidence des cas de COVID-19 ont été estimées, son impact global reste à évaluer. En effet, il a été constaté que le confinement avait fait baisser la pollution atmosphérique, ce qui aurait pu diminuer la part de la pollution dans la mortalité³⁶. A l'inverse, il est possible que le report des interventions non-urgentes à l'hôpital, ou encore l'hésitation des personnes souffrant de maladies chroniques à consulter leur médecin dans ce contexte, soit responsable, à plus ou moins brève échéance, d'une surmortalité indirectement liée au COVID-19.

34 King, *et al.* « Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola ». *Proceedings of the Royal Society B: Biological Sciences* 282, n° 1806 (7 mai 2015): 20150347. <https://doi.org/10.1098/rspb.2015.0347> ; Christley, *et al.* « “Wrong, but Useful”: Negotiating Uncertainty in Infectious Disease Modelling ». *PLoS ONE* 8, n° 10 (16 octobre 2013). <https://doi.org/10.1371/journal.pone.0076277>.

35 Fontanet, *et al.* « Cluster of COVID-19 in Northern France: A Retrospective Closed Cohort Study ». *MedRxiv*, 23 avril 2020. <https://doi.org/10.1101/2020.04.18.20071134>.

36 « En réduisant la pollution de l'air, le confinement aurait évité 11 000 décès en Europe en un mois ». *Le Monde.fr*, 29 avril 2020. https://www.lemonde.fr/planete/article/2020/04/29/en-reduisant-la-pollution-de-l-air-le-confinement-aurait-evite-11-000-deces-en-europe-en-un-mois_6038187_3244.html.

Scientifiques consultés :

Simon Cauchemez, directeur du laboratoire Modélisation mathématique des maladies infectieuses à l'Institut Pasteur, membre du conseil scientifique spécial auprès du Gouvernement.

Pascal Crépey et son équipe. Pascal Crépey est chercheur et professeur en épidémiologie et biostatistiques à l'École des hautes études en santé publique.

Samuel Alizon et son équipe. Samuel Alizon est directeur de recherche CNRS au sein du laboratoire Maladies infectieuses et vecteurs : écologie, génétique, évolution et contrôle, Université de Montpellier – IRD – CNRS.

Vittoria Colizza, directrice de recherche Inserm, et Pierre-Yves Boëlle, Professeur des universités et praticien hospitalier, responsable de l'équipe Maladies transmissibles : surveillance et modélisation à l'Institut Pierre Louis d'épidémiologie et de santé publique