

# HÉTÉROGÉNÉITÉ SOCIALE ET MODÈLES D'ÉPIDÉMIE

JEAN DOLBEAULT

RÉSUMÉ. L'épidémie de Covid-19 a mobilisé des chercheurs venus d'horizons très variés. En retour, cet effort de recherche collectif a permis de mettre en évidence de très nombreux facteurs explicatifs de l'évolution de l'épidémie. L'expérience du confinement a exposé certaines catégories de la population (les professions essentielles) plus que d'autres, et agi comme un révélateur du rôle de l'hétérogénéité sociale dans la propagation de l'épidémie. Voici quelques réflexions sur ce qu'en disent des modèles simples et sur les questions qu'ils font surgir.

Alors que les premiers résultats des enquêtes épidémiologiques sur la répartition par classes sociales, lieux d'habitation et activités professionnelles commencent à être divulgués, il est intéressant de discuter des conséquences de l'hétérogénéité du corps social sur la dynamique de l'épidémie et, très concrètement, sur la construction des modèles. Commençons par rappeler des résultats diffusés récemment par la Drees (Direction de la Recherche, des études, de l'évaluation et des statistiques) <sup>1</sup> sur l'enquête EpiCov réalisée à partir de prélèvements effectués en mai 2020, à la fin du confinement, sur un échantillon représentatif de la population en France, hors Ehpad, maisons de retraite et prisons. Selon cette enquête, la séroprévalence du SARS-CoV-2, c'est-à-dire la proportion de personnes avec des anticorps contre le virus, a été estimée à 4,5 % en France métropolitaine parmi les personnes âgées de 15 ans ou plus. La séroprévalence est la plus élevée à Paris (9,0 %), dans les départements de la petite couronne parisienne (9,5 %) et dans le Haut-Rhin (10,8 %). Les principaux facteurs de positivité au test sont le fait de vivre dans une commune à forte densité urbaine, d'exercer une profession dans le domaine du soin ou de vivre avec un nombre élevé de personnes dans le même logement. La proportion de tests positifs est également plus élevée lorsqu'un membre du ménage a présenté des symptômes ou a été testé positivement pour le SARS-CoV-2. La séroprévalence est plus élevée dans la tranche d'âge des 30-49 ans et dans le deux déciles extrêmes des niveaux de vie. Elle est également plus élevée chez les personnes immigrées nées hors de l'Europe, mais cette différence disparaît lorsqu'on prend en compte les conditions de vie et le milieu socio-économique.

---

*Date:* 18 décembre 2020.

1. Source : <https://drees.solidarites-sante.gouv.fr/IMG/pdf/er1167.pdf>

Dans les premiers temps du confinement, les statistiques publiées par Santé Publique France soulevaient une question. Alors qu'à Wuhan, le nombre de contaminations avait commencé à baisser très rapidement après l'entrée en vigueur de mesures strictes destinées à enrayer l'épidémie, en France le nombre des malades identifiés a continué à croître assez longtemps avant de se stabiliser, puis de baisser lentement. En Chine, les personnes malades ont été très rapidement isolées de leurs familles et hébergées dans des hôtels alors qu'en France ce type de mesures n'a quasiment pas été mis en place. Malgré toutes les incertitudes qui ont pesé sur les statistiques, il semblerait que l'on retrouve cette explication dans l'enquête EpiCov et que la dynamique extrêmement rapide de l'épidémie au mois de mars permette d'expliquer pour partie l'augmentation du nombre de cas après le début du confinement par des contaminations intra-familiales. La décroissance lente de l'épidémie après le maximum de la première vague, en France comme dans d'autres pays européens, nécessite toutefois d'autres explications, qui sont sans doute à rechercher du côté des professions qui n'ont pas été mises à l'arrêt ou en télé-travail. Pendant le confinement (débuté le 17 mars et levé progressivement à partir du 11 mai), certains métiers ont en effet bénéficié d'une dérogation aux mesures de confinement : les « travailleurs du soin » (personnels médicaux et paramédicaux, pharmaciens, pompiers, secouristes, ambulanciers) et les autres professions dites essentielles (aides à domicile, aides à la personne, aides ménagères, caissiers, travailleurs dans un magasin d'alimentation, livreurs à domicile, conducteurs des transports en commun, de VTC ou de taxi, responsables clientèle ou accueil des agences bancaires, salariés des stations-service, policiers, postiers, agents de nettoyage ou de propreté, agents de sécurité, artisans et salariés du bâtiment, conducteurs routiers, agriculteurs et travailleurs sociaux hors enseignants). Les taux de séroprévalence permettent aujourd'hui d'estimer tant leur exposition au virus que leur rôle dans sa propagation.

Très rapidement après le début de l'épidémie, le traçage des chaînes de contamination a mis en évidence le rôle d'individus dits super-propagateurs. Beaucoup de malades du coronavirus ne transmettent la maladie à personne, mais quelques-uns la transmettent à beaucoup. Il est aussi très clair que le risque de conséquences graves (hospitalisation, réanimation) ou à l'inverse la possibilité de ne souffrir que d'une forme bénigne, voire asymptomatique, dépend beaucoup des individus et de leurs caractéristiques, à commencer par leur âge. Il est aussi très probable que des facteurs individuels jouent sur la capacité à transmettre le virus. Dans des modèles compartimentaux simples, de type SEIR (Susceptible, Exposé, Infectieux, Retirés ou guéris), une manière de prendre en compte cette hétérogénéité consiste à introduire un taux de transmission variable. Prendre en compte l'hétérogénéité soulève plusieurs

questions : « l'une des raisons pour lesquelles cette hétérogénéité n'est souvent pas prise en compte dans les modèles est qu'il est difficile de la mesurer » souligne Gabriela Gomes dans un travail remarqué paru en mai 2020 <sup>2</sup>.

Le confinement, au mois de mars 2020, a constitué une expérience inédite. On était en droit d'attendre qu'une mesure aussi radicale ait un impact immédiat sur la transmission du virus. Pourtant, même en prenant en compte les délais d'incubation et d'aggravation des symptômes dans les cas graves, il était désespérant de voir à quel les courbes tardaient à s'infléchir. Une des raisons possibles tenait à la nature du confinement : alors que la plupart des individus voyaient leur taux de transmission drastiquement réduit, à l'inverse, ceux qui travaillaient dans les professions essentielles continuaient d'avoir une exposition au virus inchangée, voire significativement augmentée dans le cas des professions médicales. On a pu montrer <sup>3</sup> que le taux de reproduction effectif (qui mesure la croissance instantanée du nombre de malades) restait supérieur à 1 avec seulement une petite proportion d'individus dont le taux de transmission individuel n'était pas modifié par le confinement. Un tel modèle ne vaut probablement que pour les questions qu'il pose : il suggère par exemple que le taux d'immunité de groupe peut être significativement abaissé par l'hétérogénéité, de même que la taille de l'épidémie, mais que les personnes exposées professionnellement ou du fait de leur environnement socio-économique ont un taux d'individus infectés bien supérieur au reste de la population. Qualitativement, ces résultats sont tout-à-fait en ligne avec ceux de l'enquête EpiCov ; reste à en rendre compte d'un point de vue plus quantitatif, et c'est clairement d'un tout autre niveau de difficulté.

Pour faire des prédictions chiffrées ou même étudier rétrospectivement l'épidémie, on se heurte d'emblée à plusieurs difficultés. Il faut d'abord utiliser des modèles plus réalistes que de simples modèles SEIR, qu'ils soient compartimentaux ou basés sur des simulations plus ou moins aléatoires, ce qui requiert un grand nombre de mesures de natures très variées, des analyses détaillées, et en définitive la détermination d'un très grand nombre de paramètres. Quand on parle de propagation d'épidémies, il y a une notion de temps, mais aussi d'espace : on a bien vu au printemps 2020 que les différentes régions n'étaient pas touchées toutes de la même manière ou en même temps. Il y a bien sûr des études qui prennent en compte la géographie dans toute sa complexité et montrent par exemple, à l'échelle d'un pays, le rôle des autoroutes dans la propagation du virus. Comment agréger des chiffres

---

2. M. G. M. Gomes, R. M. Corder, J. G. King, K. E. Langwig, C. Souto-Maior, J. Carneiro, G. Goncalves, C. Penha-Goncalves, M. U. Ferreira, and R. Aguas. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. may 2020. <https://www.medrxiv.org/content/10.1101/2020.04.27.20081893v3>

3. J.D., G. Turinici. Heterogeneous social interactions and the COVID-19 lockdown outcome in a multi-group SEIR model. *Math. Model. Nat. Phenom.*, 15 : Paper No. 36, 18, 2020.

locaux pour donner une vue globale et pertinente de l'épidémie n'est pas aisé. Mais on se heurte aussi à la difficulté du recueil de données statistiques fiables, comparables au cours du temps et en des lieux différents. Or la remontée des informations s'est improvisée dans l'urgence, ses modalités n'ont cessé d'évoluer, quand elles n'ont pas été biaisées par des contraintes matérielles (le nombre de tests, par exemple) ou par des boucles de rétro-action complexes, au rang desquelles on peut compter les comportements de la population (les données de mobilité de Facebook au cours de la première vague de l'épidémie sont particulièrement éloquentes sur ce point).

En définitive, c'est maintenant seulement que l'on commence à avoir des données (comme celles de l'enquête EpiCov) qui permettent de caler certains paramètres des modèles, et de comprendre quels sont les mécanismes sous-jacents de la pandémie. Ces mécanismes nous apprennent peut-être autant sur la société dans laquelle nous vivons que sur la maladie induite par le coronavirus. Si l'énorme effort de recherche qui a été fait au début de l'épidémie ne se relâche pas trop tôt (comme cela a été malheureusement le cas lors des précédentes pandémies liées à d'autres virus émergents), nous en apprendrons beaucoup sur la biologie des virus, sur la gestion des pandémies et sur le fonctionnement du corps social.

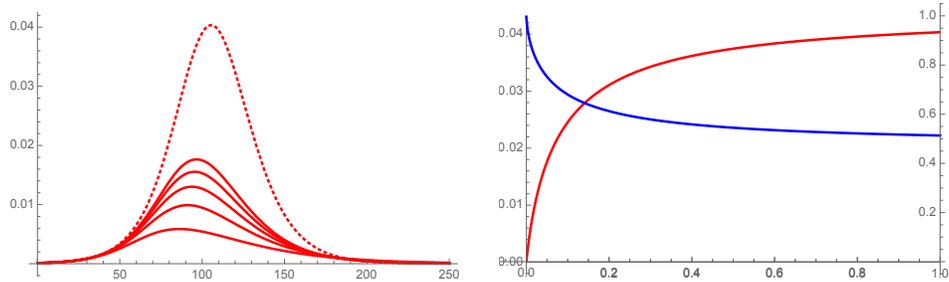


FIGURE 1. Illustration par des modèles SEIR avec un taux de reproduction de base  $R_0$  fixé à 1.37 (compatible avec les taux de la première quinzaine du confinement). A gauche, le nombre de personnes contaminées (individus infectieux ou contaminés en période d'incubation) pour une population homogène (en pointillés), et pour une population composée d'un groupe avec un taux de transmission inférieur à 1 et d'un groupe petit ( $p = 1$  à 5%) mais avec un taux de transmission élevé : pour un même  $R_0$ , le pic de l'épidémie est réduit. A droite, taille du pic de l'épidémie (en rouge) en pourcentage de la population (échelle de gauche) et pourcentage de la population indemne à la fin de l'épidémie (échelle de droite, courbe bleue) :  $p = 1$  correspond à une population homogène.