



**HAL**  
open science

# Épidémiologie mathématique & modélisation de la pandémie de COVID-19 : enjeux et diversité

Ramsès Djidjou-Demasse, Christian Selinger, Mircea T Sofonea

► **To cite this version:**

Ramsès Djidjou-Demasse, Christian Selinger, Mircea T Sofonea. Épidémiologie mathématique & modélisation de la pandémie de COVID-19 : enjeux et diversité. 2020. hal-02924338

**HAL Id: hal-02924338**

**<https://hal.science/hal-02924338v1>**

Preprint submitted on 28 Aug 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.

# Épidémiologie mathématique & modélisation de la pandémie de COVID-19 : enjeux et diversité

Ramsès Djidjou-Demasse, Christian Selinger, Mircea T. Sofonea\*  
MIVEGEC (Univ. Montpellier, CNRS, IRD), Montpellier, France  
\* auteur pour correspondance : mircea.sofonea@umontpellier.fr

## Résumé

Le champ de l'épidémiologie mathématique a connu, au cours de la pandémie de COVID-19, une production doublée d'une médiatisation exceptionnelles de ses travaux. Alors même que les données et les connaissances sur la maladie émergente étaient parcellaires, une grande diversité de modèles a été développée et appliquée dans des délais inédits, dans l'objectif d'estimer le nombre de reproduction, la date de début de l'épidémie ou l'incidence cumulée, mais aussi afin d'explorer différents scénarios d'interventions non pharmaceutiques. Leurs résultats ont largement contribué à l'épidémiosurveillance et éclairé la prise de décisions relatives aux politiques de santé publique.

**mots-clés** : épidémiologie ; modélisation ; statistiques ; nombre de reproduction ; COVID-19.

## Summary

*During the COVID-19 pandemic, the field of mathematical epidemiology experienced an exceptional production and media coverage of its work. Even though data and knowledge on the emerging disease were patchy, a wide variety of models were developed and applied in unprecedented timeframes, with the aim of estimating the reproduction number, the starting date of the epidemic or the cumulative incidence, but also to explore different scenarios of non-pharmaceutical interventions. Their results have made a major contribution to epidemiological surveillance and informed public health policy decisions.*

**keywords:** *epidemiology ; modelling ; statistics ; reproduction number ; COVID-19.*

# 1. Introduction

Fin décembre 2019, un cluster de cas de pneumonies atypiques d'étiologie inconnue est détecté à Wuhan, en Chine centrale. En l'espace d'un mois, non seulement l'agent pathogène, le coronavirus SARS-CoV-2, sera identifié et séquencé (1,2) mais les cinétiques relatives à l'histoire naturelle de la maladie respiratoire qu'il cause, telle la période d'incubation, seront elles aussi estimées (3). Dans le même temps, des approches statistiques essayant d'inférer la taille de l'épidémie à partir du nombre limité de cas détectés sont mises en ligne (4).

Deux mois plus tard, à la mi-mars 2020, alors que l'Organisation Mondiale de la Santé vient de qualifier de pandémie la propagation planétaire de la COVID-19, des simulations portant sur les leviers comportementaux de limitation de la transmission -- qualifiées d'interventions non pharmaceutiques -- explorent des scénarios d'incidence et de mortalité au Royaume-Uni avec un horizon de près de deux ans (5). L'impact scientifique, politique et médiatique de ces modèles est sans précédent dans l'histoire de l'épidémiologie des maladies infectieuses (6).

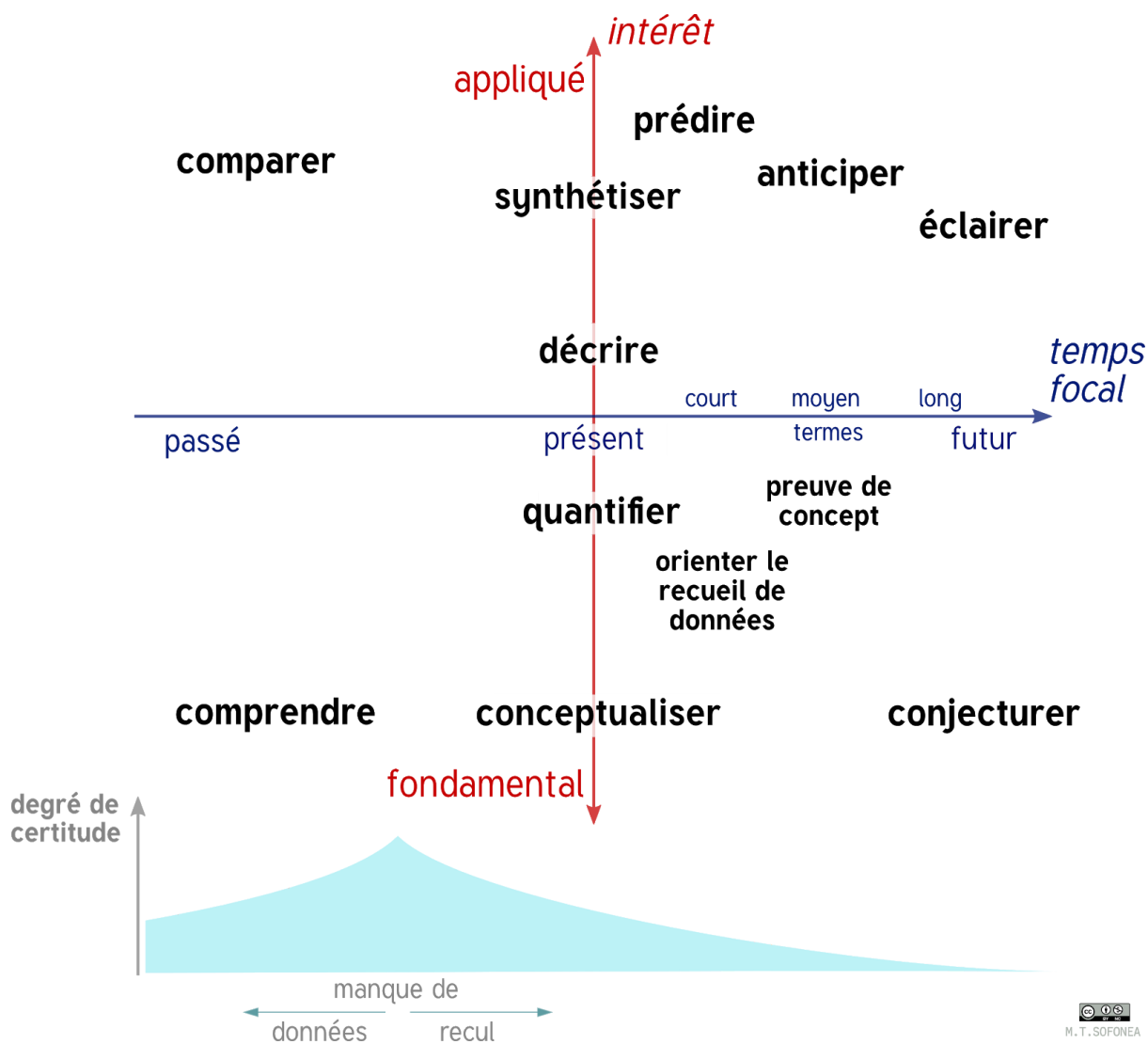
## 2. Enjeux de la modélisation

L'histoire occidentale des sciences fait habituellement remonter l'emploi de modèles mathématiques comme réponse à une question de santé publique au travail du médecin/physicien/mathématicien suisse Daniel Bernoulli (1700 -1782) sur la variolisation (7). Cette pratique importée d'Asie, où elle est attestée plusieurs siècles auparavant, consistait en l'inoculation prophylactique de la variole prélevée sur des cas peu sévères. Avec une espérance de vie à l'époque de moins de 27 ans, et une létalité de la variole estimée à 8 %, il n'était pas trivial que la variolisation systématique au jeune âge, non dénuée de risque, puisse présenter un bénéfice collectif substantiel. En utilisant une approche rigoureuse combinant la théorie, alors jeune, des équations différentielles et des données issues de la cohorte de Halley, Daniel Bernoulli estima un gain moyen de plus de trois années d'espérance de vie, apportant ainsi un argument quantitatif étayé dans un débat jusqu'alors essentiellement verbal et formulant au passage le souhait *"que dans une question qui regarde de si près le bien de l'humanité, on ne décide rien qu'avec toute la connaissance de cause qu'un peu d'analyse et de calcul peut fournir"*.

À l'image de cet exemple fondateur, l'épidémiologie mathématique contemporaine fournit un cadre de raisonnement adapté, manipulable et réfutable pour l'étude de la propagation des maladies infectieuses, là où l'intrication des différents processus (physiopathologie, transmission, politiques sanitaires) est un obstacle à l'intuition et à l'argumentation qualitative. En outre, elle permet d'apporter des éléments de réponse à des problèmes pour lesquels les approches empiriques courantes (e.g. tests in vitro, essais contrôlés) sont limitées (par le temps, l'espace, l'éthique), sinon impossibles. Tout comme ces dernières, qui y parviennent par contrainte ou randomisation, une approche par modélisation doit, pour être significative et pertinente, maximiser la part expliquée de la variance d'un système qui présente intrinsèquement de nombreux degrés de libertés. C'est pourquoi **tout modèle est**

**par essence une simplification arbitraire du phénomène étudié dont le choix est gouverné par la ou les questions auxquelles il est censé répondre.** Elles relèvent généralement d'un des trois grands objectifs suivants : *i*) la description, *ii*) la compréhension et *iii*) la prévision. Pour les illustrer dans le contexte des travaux portant sur la COVID-19, citons respectivement *i*) l'estimation de sa létalité (8), *ii*) l'existence de cas antérieurs au début de la vague épidémique (9), *iii*) la dynamique d'une hypothétique seconde vague selon différents scénarios de contrôle (10).

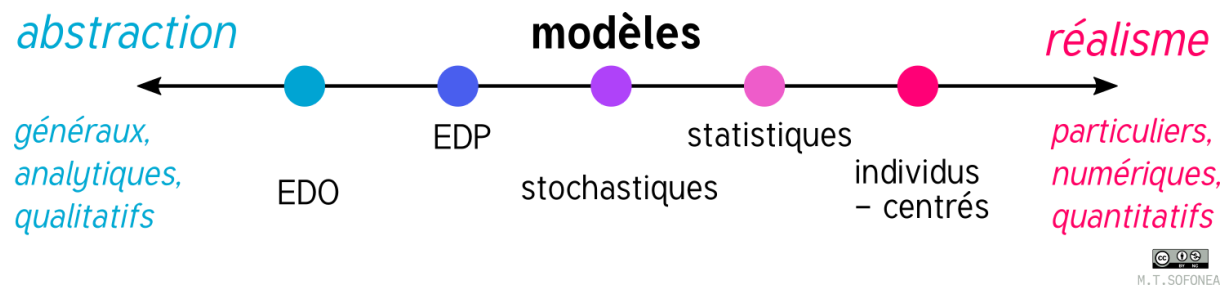
Une typologie plus détaillée des objectifs des modèles peut être caractérisée au moyen de deux axes, ainsi que représenté en Figure 1. Le premier porte sur la temporalité du phénomène par rapport à l'étude : immédiate (*nowcasting*), rétrospective ou prospective. Le second s'identifie à l'intérêt, fondamental ou appliqué, des résultats apportés par le modèle. Cette typologie recouvre une importante diversité structurale et méthodologique de modèles, dont les principales catégories peuvent être distinguées selon leur degré de réalisme, illustré en Figure 2.



**Figure 1** - Typologie des questions abordées par les modèles.

La disposition absolue des objectifs dans le repère ainsi que la distribution du degré de certitude est arbitraire et n'a qu'une valeur indicative.

Tous les aspects d'un phénomène comme la pandémie de COVID-19 ne peuvent être abordés par un seul et unique modèle. Bien qu'une modélisation puisse éventuellement répondre à plusieurs questions connexes, sa transposition à une autre temporalité, une autre échelle spatiale voire à une autre maladie infectieuse, ne peut se faire qu'au moyen d'ajustements importants qui sont au cœur de l'analyse critique de cette littérature d'interface. Dès lors, une décision relative à la santé publique ne saurait être prise sur la base d'une simulation isolée, mais devrait se reposer sur un ensemble de résultats congruents issus de modèles indépendants et, si possible, méthodologiquement complémentaires, pondérés par leur adéquation aux données présentes ou passées.



**Figure 2 - Principales catégories méthodologiques de modèles rencontrés en épidémiologie mathématiques.**

EDO : équations différentielles ordinaires. EDP : équations aux dérivées partielles.

### 3. Analyses statistiques

Toute analyse statistique est l'application d'un modèle. L'omniprésence de ces approches dans la littérature biomédicale explique peut-être pourquoi la modélisation qui sous-tend les tests d'hypothèse ou l'estimation de risques relatifs ne connaît pas la même réticence que les modèles dynamiques évoqués dans les prochaines sections. Pourtant, les régressions appliquées à des données longitudinales, reposent tout autant sur des hypothèses d'ordre mécanistique ou cinétique, comme par exemple les risques proportionnels (indépendants du temps) dans la régression de Cox classique. De la même façon, afin de s'affranchir d'hypothèses trop stringentes et potentiellement non vérifiées, des généralisations sont possibles, à condition de disposer de suffisamment de signal pour spécifier des paramètres supplémentaires.

Dans la pandémie de COVID-19, les analyses statistiques ont été décisives dans l'inférence précoce -- malgré la qualité hétérogène des données et leur nombre limité -- des paramètres qui régissent la dynamique épidémiologique : le temps d'incubation (11), l'intervalle sériel (12), le nombre de reproduction de base (3), la létalité (stratifiée en âge) (8), sont pour certains d'un intérêt clinique immédiat (p.ex. en identifiant la population pédiatrique à moindre risque de complications respiratoires) (13), mais sont aussi indispensables à la calibration des autres classes de modèles, sans quoi ces derniers seraient purement qualitatifs. Cette dépendance accorde un rôle pionnier et central à la modélisation statistique et rappelle l'enjeu majeur que représente le recueil de données exhaustives, indépendantes et non biaisées, dans les premières semaines de propagation d'une maladie infectieuse émergente.

Citons deux exemples d'analyses statistiques impliquées dans l'estimation de paramètres clés dans la pandémie de COVID-19. Tout d'abord, sa létalité (IFR, *infection fatality ratio* en anglais) stratifiée en âge : plusieurs études (8,14,15) se sont par exemple basées sur la mortalité observée parmi les personnes à bord de la croisière du *Diamond Princess* (en février 2020, au Japon), données qu'il fallait à ce moment-là corriger pour le fait que l'issue de la maladie (guérison ou décès) n'était pas encore connue pour tous les cas. L'autre exemple est le médiateur nombre de reproduction, noté  $R$ . Ce nombre moyen de cas secondaires par cas primaire, qui quantifie le potentiel de propagation de l'épidémie, a été notamment estimé par l'algorithme *EpiEstim* (16,17). Cette méthode repose sur la maximisation de la probabilité d'observer autant de nouveaux cas connaissant l'incidence passée, sous l'hypothèse que  $R$  soit resté constant sur la période estimée. Sa définition prospective a l'avantage de pouvoir détecter les changements récents dans le régime de transmission d'une épidémie encore en cours (pour laquelle le nombre de cas causés par la dernière cohorte d'infectés n'est pas encore connu).

Fondamentalement probabilistes, les modèles statistiques permettent d'estimer une valeur moyenne extraite à partir des données qui ne la mesure pas directement et conjointement d'en quantifier l'incertitude (par des intervalles de confiance, des tests d'hypothèses), voire d'inférer une distribution qui en cerne la variabilité. Cependant, la capacité inférentielle de ces approches est réduite aux paramètres dont les variations sont relativement bien

reflétées dans les données. Ainsi, pour reconstruire initialement la trajectoire épidémique de la COVID-19 en population en France à partir des seules données hospitalières, l'approche statistique était insuffisante car un trop grand nombre d'événements inconnus s'intercalent entre les événements (largement majoritaires) de contaminations entre cas non sévères et le flux d'hospitalisations observé deux semaines plus tard. Prennent alors le relai les modèles explicitement dynamiques, qui formalisent le mécanisme de propagation ainsi que la cinétique de l'histoire naturelle de la maladie et son éventuelle prise en charge.

## 4. Modèles mécanistiques déterministes

L'étude de la dynamique épidémique de la COVID-19 nécessite de faire appel à une ou plusieurs variables structurantes : le temps absolu, l'âge des hôtes, l'âge de leurs infections, leur statut immunologique, etc. Le temps est essentiel puisqu'il sous-tend la description des variations continues de ces systèmes appelés dynamiques, majoritairement formalisés sous forme d'équations différentielles ordinaires (EDO) (18). Plus rarement, le temps peut être considéré discontinu (à l'instant des données quotidiennes), le système prenant alors la forme de suites récurrentes (10).

L'un des processus clés dans la propagation d'une épidémie est la force d'infection (la vitesse à laquelle un hôte sensible devient infecté). Dans sa modélisation la plus simple dite *SIR* de Kermack et McKendrick (19), la population d'hôte est répartie en trois compartiments : les individus sains susceptibles d'être infectés *S*, les individus infectés contagieux *I* et les retirés *R* (guéris immunisés ou décédés). Hors maladie vectorielle, on suppose que l'infection se propage par contact direct entre individus sains du groupe *S* et individus infectés du groupe *I*, et que l'incidence résultante est proportionnelle au produit des deux densités : c'est la loi d'action de masse. Ne nécessitant que deux paramètres, cette approche a par s'ajustent avec succès sur des données d'incidence telles que celles de la peste à Bombay entre 1905 et 1906 (19).

La contagiosité d'un individu atteint du SARS-CoV-2 affecte la propagation épidémique et varie en fonction d'un autre âge, celui de l'infection, c'est-à-dire du temps écoulé depuis que cet individu est infecté (20). Cette variable permet de prendre en compte la "mémoire" d'une infection chez une personne infectée, par opposition aux modèles dits markoviens (comme celui de Bernoulli et une grande partie des modèles encore produits actuellement), dont la probabilité par unité de temps de survenue d'un événement (i.e. un taux, au sens strict du terme), par exemple une guérison, est indépendante du temps déjà écoulé. Dans le cas de l'épidémie de COVID-19 en France, la prise en compte de cet effet a notamment permis de capturer relativement tôt la dynamique piquée de la première vague (10), un aspect qui peut s'avérer opportun dans la gestion des capacités hospitalières. Cette mémoire de l'infection peut être prise en compte suivant deux formalismes. Le premier est un formalisme discret qui consiste à stratifier la progression de l'infection chez un individu en plusieurs compartiments successifs (10,15). Le second est un formalisme continu qui consiste à considérer l'âge de l'infection comme une variable structurante continue (21), le système dynamique est alors régi par des équations aux dérivées partielles (EDP).



L'apparition du SARS-CoV-2 dans une population sans immunité préexistante a conduit de nombreux pays à appliquer des mesures sanitaires (confinement, port du masque, dépistage et traçage de contact) afin de contenir l'épidémie le temps du déploiement d'une solution pharmaceutique. Les premiers modèles mécanistiques ont exploré les meilleures stratégies d'implémentation de contrôle sanitaire (22,23). Ces modèles faisaient l'hypothèse d'une population homogène -- p.ex., sans prise en compte d'une structure d'âge de la population -- et ne prenaient pas explicitement en compte la mémoire de l'infection comme dans certains modèles développés par la suite (10,15,21).

Les modèles dynamiques déterministes, qu'ils reposent sur des équations différentielles ou aux dérivées partielles, sont faciles à simuler. Leur souplesse autorise l'exploration d'une grande diversité de scénarios, tandis que de nombreux outils théoriques et numériques permettent de les exploiter à des fins analytiques (trouver une formule permettant d'exprimer le nombre de reproduction en fonction des autres paramètres en présence) et statistiques (ajustement de paramètres, analyse de sensibilité). Plus généralement, ils constituent un moyen parcimonieux (donc calibrable avec un minimum de données) pour appréhender le comportement non trivial des trajectoires épidémiques, lesquelles résultent d'interactions non linéaires entre les densités des différents compartiments clinico-épidémiologiques. Malgré leur simplicité structurale, les travaux de méta-modélisation suggèrent que ces modèles, qui moyennent implicitement de nombreux aspects du phénomène (taux de transition entre compartiments, distribution spatiale), génèrent des résultats robustes et conservatifs d'un point de vue de santé publique, confortant leur utilité, ne serait-ce qu'en début d'épidémie (24).

## Encadré technique : deux expressions remarquables du nombre de reproduction

Le nombre de reproduction  $R$  représente le nombre moyen d'infectés secondaires qu'un cas contamine au cours de sa période de contagiosité. Il est souvent dit 'de base' et désigné par  $R_0$  en début d'épidémie, en absence de mesure sanitaire et lorsque l'ensemble de la population est immunologiquement naïve pour la maladie considérée. L'épidémie est en croissance si et seulement si  $R$  est supérieur à 1.

Dans le cas général, l'estimation de  $R$  au jour  $t$  se fait par le ratio de l'incidence du jour,  $c_t$ , par les incidences des jours précédents,  $c_{t-s}$ , pondérés rétrospectivement par leur intervalle sériel  $w_s$ . Ce dernier peut être vue comme la contribution relative de chaque jour post-infection au potentiel total de contagiosité sur toute la durée infectieuse, soit

$$R_t = \frac{c_t}{\sum_{s>0} c_{t-s} w_s}.$$

Cette expression est particulièrement pratique en épidémiologie dans la mesure où elle ne nécessite pas de modéliser explicitement l'histoire naturelle de la maladie et s'applique aussi lorsque la trajectoire épidémique n'est pas exponentielle (p.ex. lorsque de nouvelles mesures sanitaires sont mises en place). En revanche, cet estimateur est sensible aux variations de l'effort de détection des cas sur la période considérée.

Une autre relation remarquable impliquant le nombre de reproduction est issue du modèle classique  $SIR$ , qui suppose une durée de contagiosité distribuée exponentiellement et de moyenne  $D$ . Si  $T$  est le temps de doublement, c'est-à-dire la durée au bout de laquelle l'incidence double, alors

$$(R - 1)T = D \ln 2.$$

Cette relation formalise le fait qu'à durée de contagiosité constante et connue, le nombre de reproduction permet aussi d'estimer la cinétique de l'épidémie. On notera en particulier que la relation inversement proportionnelle entre l'écart à 1 du nombre de reproduction et le temps de doublement implique que l'arrivée d'un pic épidémique peut être considérablement repoussée par une baisse modeste du nombre de contacts, pour les maladies infectieuses de nombre de reproduction proche de 1. À l'inverse, un léger relâchement collectif de mesures préventives ayant maintenu jusqu'alors une épidémie sous contrôle peut engendrer un rebond rapide. Le cas de la circulation de la COVID-19 en France au cours de l'été 2020 est une illustration de cet effet.

## 5. Approches stochastiques

Le processus de contagion est de nature aléatoire à toutes les échelles. De la liaison d'une particule virale à la surface d'une cellule humaine au conditionnement de l'information génétique répliquée, de l'excrétion de nouveaux virions à la transmission d'un second individu, du comportement individuel à la mobilité dans une population, tous ces facteurs privilégiant l'émergence d'une épidémie sont le fruit d'une longue chaîne d'événements aléatoires, de probabilités variables et plus ou moins indépendants de la survenue des autres. Les modèles mathématiques en épidémiologie se sont pendant longtemps intéressés davantage à la valeur moyenne de ces facteurs, négligeant leurs fluctuations. Pourtant, l'approche stochastique offre deux avantages majeurs. D'abord, elle permet d'intégrer les informations relatives à la variabilité des paramètres utilisés et aux corrélations statistiques observés entre eux. Deuxièmement, le calcul de probabilités permet de quantifier aussi le début et la fin d'une épidémie en terme de probabilité d'émergence et d'extinction et de la distribution de la taille finale. Comme le trait d'union entre l'approche probabiliste et déterministe est la loi des grands nombres, un tel modèle peut facilement être conçu pour que l'évolution moyenne d'une épidémie corresponde à celle d'un modèle déterministe.

Ainsi, un modèle stochastique avec un taux de transmission aléatoire (25) a pu mettre en évidence la nécessité de plusieurs introductions indépendantes pour expliquer le démarrage de l'épidémie COVID-19 à Wuhan. Prenant en compte l'aléatoire dans les réseaux de contacts et l'hétérogénéité du nombre d'infections secondaires, un autre modèle (26) a démontré que l'estimation de la taille finale de l'épidémie est sujette à une incertitude importante, surtout quand les données disponibles portent seulement sur les premières semaines de l'épidémie.

Parmi les modèles stochastiques, les processus de branchement occupent une place importante dans l'application à l'épidémiologie. Partant de l'idée qu'une nouvelle infection est un événement aléatoire distribué de façon identique et indépendante, on suppose que chaque individu infectieux peut produire de nouvelles infections selon cette loi. Toute réalisation d'un tel processus conduit à un arbre, et un branchement correspond à des nouvelles infections engendrées par sa jonction. Plusieurs modèles basés sur les processus de branchement ont été mis en avant pour étudier l'efficacité du traçage et de l'isolement au début de l'épidémie, quand l'évènement de branchement n'est pas limité par un manque d'individus susceptibles (27). En particulier, ces modèles suggèrent que le traçage électronique nécessiterait une fraction importante d'utilisateurs pour que l'épidémie devient sous-critique.

L'approche stochastique est particulièrement indiquée lorsqu'un paramètre clé du processus infectieux (tel le nombre de reproduction individuel) présente une forte variance parmi les cas et que l'on souhaite quantifier l'impact de cette hétérogénéité sur la dynamique épidémiologique, en particulier dans les contextes où les effectifs sont faibles (initiation ou extinction de l'épidémie, propagation locale).

## 6. Modèles individus-centrés

Le développement de modèles à compartiments (systèmes dynamiques) préconise une sous-division de la population étudiée par catégorie de contribution épidémiologique équivalente (selon le statut infectieux, selon l'âge, le sexe...). L'intégration multi-échelle d'aspects démographiques, géographiques, intra-hôte ou comportementaux de l'épidémie de COVID-19 peut rendre ces outils mathématiques inapplicables. Les modèles individu-centrés (MIC) permettent de dépasser ces limites en utilisant des approches informatiques de calcul de haute performance. Ainsi, chaque individu au sein de la population est considéré comme un objet avec un certain nombre d'attributs déterminant de façon individuelle la contagiosité, la susceptibilité mais aussi les interactions avec d'autres individus. En intégrant des données de densité de population de haute résolution, les MIC ont été utilisés pour évaluer l'impact de stratégies nationales de suppression ou d'atténuation au début de l'épidémie de COVID-19 (28).

Pour la France, un MIC (29) a mis en avant l'importance de l'isolement des personnes à haut risque, et le MIC *epidemap*<sup>1</sup> a pu anticiper la disparité géographique de cas critiques grâce à un modèle détaillé de mobilité humaine. Tout en étant très flexible dans sa conception, le temps de calcul d'un MIC peut devenir un facteur limitant si le nombre d'individus simulés est trop grand ou si les mécanismes d'interactions individuelles sont trop complexes. Ainsi, d'autres MIC ont été proposés pour évaluer en continu l'impact épidémiologique et économique de différentes combinaisons d'interventions non-pharmaceutiques à l'échelle d'une zone métropolitaine (30–34).

Complémentaire aux modèles compartimentaux, les MICs constituent une approche de choix pour améliorer notre connaissance et générer de nouvelles hypothèses quant aux mécanismes spatiaux et comportementaux de l'épidémie de COVID-19. Ils permettent en outre d'explorer des scénarios réalistes sur le plan géographique et sont appelés à devenir de précieux outils dans les échelons territoriaux de l'épidémiosurveillance et de la gestion sanitaire.

## 7. Conclusion

La modélisation en épidémiologie des maladies infectieuses repose aujourd'hui sur une diversité de méthodes, fortes de décennies de développements mathématiques et informatiques, à l'interface avec les sciences du vivant. La multiplication d'approches complémentaires -- statistiques, analytiques et numériques -- d'un même sujet permet de résoudre des questions inaccessibles par des approches intuitives ou expérimentales, à l'image des travaux rapidement réalisés dès le début de la pandémie de COVID-19, qui ont fourni des outils précieux pour l'épidémiosurveillance et l'anticipation sanitaire dans un contexte de données limitées. Les modèles doivent être constamment enrichis par des données de terrain. Leur pertinence doit être évaluée à la lueur de l'adéquation à ces dernières et de la validité des hypothèses sous-jacentes. Les modèles sont en particulier au

---

<sup>1</sup> [epidemap.com](http://epidemap.com)

cœur des analyses statistiques qui permettent de distinguer le bruit des signaux -- et comparer ces derniers entre eux --, un enjeu méthodologique majeur dans les études de santé publique.

### **Points à retenir**

1. La modélisation fournit un cadre de raisonnement formel, manipulable et réfutable pour l'étude de la propagation des maladies infectieuses.
2. Les enjeux de la modélisation résident dans la compréhension des épidémies passées, la quantification d'une dynamique en cours, et l'anticipation de la trajectoire à venir.
3. La modélisation statistique constitue la pierre angulaire de l'épidémiologie mathématique d'une maladie infectieuse émergente, sur lesquelles reposent les autres méthodes (dynamiques déterministes, probabilistes, individus-centrés).
4. Le dialogue permanent entre les modèles et le terrain stimule la recherche en maladies infectieuses en orientant le recueil de données et en suggérant de nouvelles hypothèses à tester.
5. La modélisation permet d'explorer des scénarios inaccessibles à l'expérience et peut ainsi éclairer la prise de décision en santé publique, sous réserve d'une multiplicité de résultats indépendants compatibles avec les données empiriques.

### **Déclaration d'intérêts**

Les auteurs déclarent ne pas avoir de conflits d'intérêts en relation avec cet article.

### **Remerciements**

Les auteurs remercient Yannis Michalakis, Samuel Alizon, Marc Choisy pour leurs relectures attentives et leurs commentaires, et l'ensemble de l'équipe de modélisation ETE pour les discussions partagées.

## Références

1. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. mars 2020;579(7798):265-9.
2. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. The species Severe acute respiratory syndrome-related coronavirus : classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. avr 2020;5(4):536-44.
3. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus - Infected Pneumonia. *N Engl J Med*. 26 mars 2020;382(13):1199-207.
4. Imai N, Dorigatti I, Cori A, Riley S, Ferguson N. Estimating the potential total number of novel Coronavirus cases in Wuhan City, China. [imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/](http://imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/); 2020 janv. Disponible sur: <http://spiral.imperial.ac.uk/handle/10044/1/77149>
5. Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID- 19 mortality and healthcare demand. [imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/](http://imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/); 2020 p. 20. Disponible sur: <http://hdl.handle.net/10044/1/77482>
6. Adam D. Special report: The simulations driving the world's response to COVID-19. *Nature*. 2 avr 2020;580(7803):316-8.
7. Bernoulli D. Essai d'une nouvelle analyse de la mortalité causée par la petite Vérole, et des avantages de l'Inoculation pour la prévenir. *Hist Académie R Sci Avec Mém Mathématique Phys*. 1760;1-45.
8. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*. mars 2020;S1473309920302437.
9. Althouse BM, Wenger EA, Miller JC, Scarpino SV, Allard A, Hébert-Dufresne L, et al. Stochasticity and heterogeneity in the transmission dynamics of SARS-CoV-2. *ArXiv200513689 Phys Q-Bio*. 27 mai 2020; Disponible sur: <http://arxiv.org/abs/2005.13689>
10. Sofonea MT, Reyné B, Elie B, Djidjou-Demasse R, Selinger C, Michalakakis Y, et al. Epidemiological monitoring and control perspectives: application of a parsimonious modelling framework to the COVID-19 dynamics in France. *medRxiv*. 24 mai 2020;2020.05.22.20110593.
11. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung S, et al. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. *J Clin Med*. févr 2020;9(2):538.
12. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis*. avr 2020;93:284-6.
13. Phua J, Weng L, Ling L, Egi M, Lim C-M, Divatia JV, et al. Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations. *Lancet Respir Med*. avr 2020;S2213260020301612.
14. Russell TW, Hellewell J, Jarvis CI, van-Zandvoort K, Abbott S, Ratnayake R, et al. Estimating the infection and case fatality ratio for COVID-19 using age-adjusted data from the outbreak on the Diamond Princess cruise ship [Internet]. *Epidemiology*; 2020 mars. Disponible sur: <http://medrxiv.org/lookup/doi/10.1101/2020.03.05.20031773>
15. Salje H, Kiem CT, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. *Science* [Internet]. 13 mai 2020; Disponible sur:

- <https://science.sciencemag.org/content/early/2020/05/12/science.abc3517>
16. Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *Am J Epidemiol.* 1 nov 2013;178(9):1505-12.
  17. Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics.* 1 déc 2019;29:100356.
  18. Keeling MJ, Rohani P. *Modeling infectious diseases in humans and animals.* Princeton University Press; 2008.
  19. Kermack W, McKendrick A. Contributions to the mathematical theory of epidemics, part 1. *Proc R Soc Math Phys Eng Sci.* 1927;
  20. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science.* 31 mars 2020; Disponible sur: <https://science.sciencemag.org/content/early/2020/03/30/science.abb6936>
  21. Richard Q, Alizon S, Choisy M, Sofonea MT, Djidjou-Demasse R. Age-structured non-pharmaceutical interventions for optimal control of COVID-19 epidemic. *medRxiv.* 23 juin 2020;2020.06.23.20138099.
  22. Djidjou-Demasse R, Michalakis Y, Choisy M, Sofonea MT, Alizon S. Optimal COVID-19 epidemic control until vaccine deployment [Internet]. *Infectious Diseases (except HIV/AIDS);* 2020 avr. Disponible sur: <http://medrxiv.org/lookup/doi/10.1101/2020.04.02.20049189>
  23. Perkins A, Espana G. Optimal control of the COVID-19 pandemic with non-pharmaceutical interventions. *medRxiv.* 23 avr 2020;2020.04.22.20076018.
  24. Trapman P, Ball F, Dhersin J-S, Tran VC, Wallinga J, Britton T. Inferring  $R_0$  in emerging epidemics—the effect of common population structure is small. *J R Soc Interface.* 31 août 2016;13(121):20160288.
  25. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Dis.* 11 mars 2020; Disponible sur: <http://www.sciencedirect.com/science/article/pii/S1473309920301444>
  26. Hébert-Dufresne L, Althouse BM, Scarpino SV, Allard A. Beyond  $R_0$ : Heterogeneity in secondary infections and probabilistic epidemic forecasting. *medRxiv.* 16 avr 2020;2020.02.10.20021725.
  27. Levesque J, Maybury DW, Shaw RD. A model of COVID-19 propagation based on a gamma subordinated negative binomial branching process: A tool for decision making with small populations. *medRxiv.* 10 juill 2020;2020.07.08.20149039.
  28. Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. 2020 mars. Disponible sur: <http://spiral.imperial.ac.uk/handle/10044/1/77482>
  29. Hoertel N, Blachier M, Blanco C, Olfson M, Massetti M, Rico MS, et al. A stochastic agent-based model of the SARS-CoV-2 epidemic in France. *Nat Med.* 14 juill 2020;1-5.
  30. Covasim: an agent-based model of COVID-19 dynamics and interventions. *medRxiv.* Disponible sur: <https://www.medrxiv.org/content/10.1101/2020.05.10.20097469v1>
  31. An agent-based epidemic model REINA for COVID-19 to identify destructive policies | *medRxiv.* Disponible sur: <https://www.medrxiv.org/content/10.1101/2020.04.09.20047498v2>
  32. Silva P, Batista P, Seixas Lima H, Alves M, Pedrosa Silva R, Guimarães F. COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. *Vol. 139.* 2020.

33. Drogoul A, Taillandier P, Gaudou B, Choisy M, Chapuis K, Huynh QN, et al. Designing social simulation to (seriously) support decision-making: COMOKIT, an agent-based modelling toolkit to analyse and compare the impacts of public health interventions against COVID-19.
34. Bossert A, Kersting M, Timme M, Schröder M, Feki A, Coetzee J, et al. Limited containment options of COVID-19 outbreak revealed by regional agent-based simulations for South Africa. ArXiv200405513 Phys Q-Bio. 11 avr 2020. Disponible sur: <http://arxiv.org/abs/2004.05513>