

Pourquoi étudier les maladies infectieuses quand on est un biologiste de l'évolution ?

Il existe beaucoup de liens entre l'écologie évolutive et l'épidémiologie. D'abord, il y a une grande similitude méthodologique et conceptuelle. Les outils utilisés en épidémiologie sont, au départ, des outils de la dynamique des populations servant à suivre des fluctuations dans les tailles de populations, qu'il s'agisse de mammifères ou de bactéries.

Il y a aussi des liens plus biologiques. Les maladies infectieuses sont principalement causées par des microbes – virus, bactéries, parasites – qui évoluent très vite. En vingt ans, on dénombre à peu près autant de générations du virus de l'immunodéficience humaine (VIH) au sein d'une personne porteuse que de générations d'*Homo sapiens* depuis son apparition. De plus, la vitesse d'évolution des microbes est plus élevée que dans d'autres circonstances, car ils affrontent des contraintes sélectives très fortes du fait des réponses immunitaires de leurs hôtes.

Cette évolution rapide peut modifier la manière dont les épidémies se propagent. On parle de « boucle de rétroaction ». La propagation des épidémies façonne des pressions de sélection sur les microbes et, en retour, leur évolution modifie la façon dont ils se propagent. Prenez l'antibiorésistance. La propagation d'une bactérie dans la population entraîne l'utilisation de traitements antibiotiques, lesquels engendrent une pression de sélection qui favorise la propagation de souches bactériennes antibiotiques. D'où l'importance d'étudier en même temps la propagation d'une épidémie et l'évolution du microbe qui en est responsable.

La modélisation des épidémies s'est-elle toujours appuyée sur cette approche évolutive ?

Non, il y a eu un va-et-vient des outils théoriques entre la dynamique des populations et l'épidémiologie. Une personne qui a beaucoup contribué à renforcer ce lien est Robert May. Ce physicien théoricien britannique d'origine australienne s'est pris de passion pour l'écologie théorique dans les années 1970. Il a apporté des contributions majeures à la dynamique des populations animales avant de contribuer à l'essor de l'épidémiologie mathématique moderne. Dans son étude sur le virus de la myxomatose en 1982, il a aussi été parmi les premiers à inclure l'évolution biologique en épidémiologie. Mais cela est vraiment rentré dans les mœurs dix à vingt ans plus tard.

En quoi consiste la modélisation des épidémies ?

Il y a principalement trois grandes raisons de modéliser. Souvent, le modélisateur se focalise sur l'une des trois, ce qui influe sur la démarche et les outils utilisés. La première raison est la description. À partir du moment où

l'on suit une épidémie, où l'on représente, par exemple, une carte de France avec le nombre de cas, un modèle est sous-jacent. Cette partie descriptive met en jeu principalement des outils de la modélisation statistique. On essaye de décrire au mieux l'épidémie au vu des observations et en tenant compte du fait qu'une observation est toujours biaisée – on n'observe jamais le virus en train de se transmettre, on ne voit que des manifestations de l'épidémie.

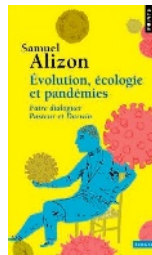
La deuxième raison de modéliser est la compréhension. Un modèle commence d'ailleurs toujours par présenter ses hypothèses. Cela permet de dialoguer, de confronter les idées et les interprétations. Pour ce type d'approches, on s'appuie en général sur une modélisation plus mathématique. En pratique, on utilise souvent les connaissances issues d'une approche descriptive (et de modèles statistiques) pour façonner les équations reflétant les hypothèses de base.

Cette partie est très utile pour s'affranchir des biais du cerveau. Par exemple, ce dernier a du mal à appréhender la croissance exponentielle, comme celle d'une épidémie qui doublerait en taille tous les trois jours. Le fait qu'il faille la même durée pour passer de 16 à 32 personnes infectées et de 160 000 à 320 000 est complètement contre-intuitif. Avec un modèle, cela devient trivial. De même, notre cerveau perçoit mal les interactions de phénomènes, comme la variation de la létalité de l'infection en fonction de l'âge, un ingrédient clé pour comprendre la dynamique hospitalière dans le cadre du Covid-19.

Ces modèles mathématiques aident à décorréler les processus sous-jacents, à examiner sous quelles conditions on reproduit les tendances observées. Il y a donc toujours une question de recherche associée et, selon la question posée, on n'utilisera pas le même modèle. En fait, pour une même question on pourra élaborer des modèles différents, du fait de leurs hypothèses ou des techniques qu'ils mettent en jeu. C'est pourquoi il y a une vraie valeur ajoutée à avoir différentes équipes confrontant leurs approches. Si les analyses concordent, ce consensus donne plus de poids à l'interprétation. Si elles divergent, c'est une motivation pour mieux comprendre les processus.

On modélise aussi pour anticiper...

En effet, la prédiction est la troisième raison de modéliser même si on préfère parler de « scénarios », car personne ne peut prédire l'avenir. La propagation d'une épidémie dépend de trois composantes. La première est biologique. Elle inclut des paramètres classiques tels que la contagiosité ou la durée de l'infection, mais aussi des phénomènes plus variables tels que l'évolution du virus, de sa virulence, de son échappement au système immunitaire. Au



À LIRE

Samuel Alizon a récemment publié **Évolution, écologie et pandémies**, Seuil, 2020.

début de l'épidémie de Covid-19, cette composante était une inconnue majeure. En février 2020, on ignorait, par exemple, la proportion des personnes infectées qui décédaient, une information essentielle pour élaborer le moindre modèle.

La deuxième composante est physique : comment l'épidémie se propage une fois que l'on a pris en compte les contraintes biologiques. C'est la partie la plus « facile », parce que les lois de la physique s'appliquent assez bien au processus de contagion. Mais une troisième composante vient contrecarrer les prédictions : la composante sociale. Les comportements peuvent changer du tout au tout d'un jour à l'autre. Si demain le gouvernement confine le pays, ou si, à l'inverse, il relâche tous les gestes barrières, l'épidémie suivra des trajectoires diamétralement opposées. Prévoir l'état de l'épidémie à plus de deux semaines nécessite donc de prévoir les politiques de santé...

C'est pourquoi on explore différents scénarios. Celui de référence est en général la tendance actuelle. Puis on explore des scénarios qui dévient de cette référence. Nous sommes début septembre : que se passe-t-il si la rentrée a un effet que l'on ne voit pas encore à cause des délais entre l'infection et l'admission en réanimation ?

On peut aussi produire des scénarios rétrospectifs. Notre équipe l'a fait dans deux cas : un scénario regardant ce qui se serait passé si l'on n'avait pas eu de vaccin ou si les vaccins avaient uniquement empêché les formes sévères, sans empêcher la transmission, et un scénario où le troisième confinement, qui a eu lieu début avril 2021, avait commencé deux mois plus tôt, suivant les recommandations du Conseil scientifique. Là encore, on ne modélise pas exactement ce qui se serait passé, de même qu'on ne modélise pas le futur. On explore une tendance en fonction d'hypothèses données.

Utilisez-vous les scénarios rétrospectifs pour vérifier vos prévisions ?

Oui, notre équipe l'a fait sur nos scénarios prospectifs publiés depuis douze mois afin d'évaluer leur écart avec la dynamique hospitalière observée. Nous avons trouvé qu'en moyenne, notre modèle et nos hypothèses donnaient des tendances robustes sur cinq semaines, ce qui est assez satisfaisant sachant qu'à plus de deux semaines vous êtes à la merci d'un changement de politique sanitaire.

Nous avons aussi identifié pourquoi nos projections étaient trop pessimistes fin avril 2021, lors du troisième déconfinement. Le modèle fonctionnait, mais le problème venait des paramètres d'entrée, car, avec les données de l'époque, on supposait que l'épidémie décroissait moins vite qu'elle ne le faisait en réalité. Pourquoi ? On ne peut que faire des

conjectures, car, en France, on ne suit pas d'assez près l'épidémie pour avoir une réponse précise. Une hypothèse est que le troisième confinement a coïncidé avec le début des vacances scolaires et qu'une mesure de contrôle portant sur les écoles met plus de temps à se manifester. En effet, non seulement les chiffres hospitaliers ont un retard de deux semaines sur l'épidémie, mais comme les enfants font très peu de formes sévères, le décalage est encore plus important. D'ailleurs, deux semaines plus tard, le 16 mai, avec les chiffres mis à jour, le scénario a reproduit la tendance observée jusqu'à début juillet.

Quel type de données utilisez-vous dans vos modèles ?

Nous utilisons principalement les données hospitalières, pour deux raisons. D'une part, l'échantillonnage est exhaustif : toutes les personnes hospitalisées à cause du Covid-19 sont comptabilisées, alors qu'on dépiste moins de la moitié des personnes infectées. D'autre part, les admissions en réanimation ont un délai de « seulement » deux semaines entre l'infection et l'observation, alors qu'il est d'un mois avec les décès. Certes, en théorie, les dépistages permettent un suivi au plus près de l'épidémie,



Des informations sur la mobilité aident à anticiper le nombre d'hospitalisations

car les symptômes apparaissent généralement quatre jours après infection. Mais pour s'affranchir des biais, il faut une réflexion statistique. Par exemple, le Royaume-Uni organise des campagnes de dépistage aléatoires où des milliers de personnes, avec ou sans symptômes, sont testées chaque jour, avec un plan d'échantillonnage réfléchi. Ainsi, même avec moins de tests qu'en France, on a une photographie bien plus précise de l'épidémie.

Mais nous avons aussi développé d'autres modèles avec des données plus originales, comme celles de Facebook. L'idée est d'utiliser ce qu'on appelle des « signaux précoces » pour suivre l'épidémie au plus près à l'aide de modèles statistiques : connaissant la dynamique hospitalière passée, avoir des